Some remarks about conservation for residual distribution schemes

Abgrall, Rémi

Abstract: We are interested in the discretisation of the steady version of hyperbolic problems. We first show that all the known schemes (up to our knowledge) can be rephrased in a common framework. Using this framework, we then show they flux formulation, with an explicit construction of the flux, and thus are locally conservative. This is well known for the finite volume schemes or the discontinuous Galerkin ones, much less known for the continuous finite element methods. We also show that Tadmor's entropy stability formulation can naturally be rephrased in this framework as an additional conservation relation discretisation, and using this, we show some connections with the recent papers [13, 20, 18, 19]. This contribution is an enhanced version of [4].

DOI: https://doi.org/10.1515/cmam-2017-0056

Some remarks about conservation for residual distribution schemes

R. Abgrall.

November 28, 2017

Abstract

We are interested in the discretisation of the steady version of hyperbolic problems. We first show that all the known schemes (up to our knowledge) can be rephrased in a common framework. Using this framework, we then show they flux formulation, with an explicit construction of the flux, and thus are locally conservative. This is well known for the finite volume schemes or the discontinuous Galerkin ones, much less known for the continuous finite element methods. We also show that Tadmor's entropy stability formulation can naturally be rephrased in this framework as an additional conservation relation discretisation, and using this, we show some connections with the recent papers [1, 2, 3, 4]. This contribution is an enhanced version of [5].

1 Introduction

In this paper, we are interested in the approximation of non-linear hyperbolic problems. To make things more precise, our target are the Euler equations in the compressible regime, other examples are the MHD equations. The case of parabolic problems in which the elliptic terms play an important role only in some area of the computational domain, such as the Navier-Stokes equations in the compressible regime, or the resistive MHD equations, can be dealt with in a similar way. In a series of papers [6, 7, 8, 9, 10, 11, 12, 13, 14, 15], following the pioneering work of Roe and Deconinck [16], we have developed, with collaborators[1], a class of schemes that borrow some features from the finite element methods, and others, such as a local maximum principle and a non-linear stabilisation from the finite difference/finite volume methods. Though the methods have been developed with some rigour, there is a lack of a more theoretical analysis, and also to explain in a clearer way the connections with more familiar methods such as the continuous finite elements methods or the discontinuous Galerkin ones.

The ambition of this paper is to provide this link through a discussion about conservation and entropy stability. In most of the paper, we consider steady problems in the scalar case. The extension to the system case is immediate. Examples of schemes are given in the paper and the appendix. Their extensions to the system case can be found in [11] for the pure hyperbolic case and in [13, 14] for the Navier Stokes equations.

The model problem is

$$\text{div } \mathbf{f}(u) = 0 \qquad \text{in } \Omega \tag{1a}$$

subjected to

$$\min(\nabla_u \mathbf{f}(u)\cdot \mathbf{n}(\mathbf{x}), 0)(u-u_b) = 0 \text{ on } \partial\Omega. \tag{1b}$$

The domain Ω is assumed to be bounded, and regular. We assume for simplicity that its boundary is never characteristic. We also assume that it has a polygonal shape and thus any triangulation that we consider covers Ω exactly. In (1b), $\mathbf{n}(\mathbf{x})$ is the outward unit vector at $\mathbf{x}\in\partial\Omega$ and u_b is a regular enough function. The weak formulation of (1) is: $u\in L^\infty(\Omega)$ is a weak solution of (1) if for any $\varphi\in C_0^1(\Omega)$,

$$-\int_\Omega \nabla v\cdot \mathbf{f}(u^h)\, d\mathbf{x} + \int_{\partial\Omega} v\big(\mathcal{F}_{\mathbf{n}}(u,u_b) - \mathbf{f}(u)\cdot\mathbf{n}\big)\, d\gamma = 0 \tag{2}$$

[1] in particular M. Ricchiuto, from INRIA Bordeaux Sud-Ouest

where $\mathcal{F}_{\mathbf{n}}$ is a flux that is almost everywhere the upwind flux:

$$\mathcal{F}_{\mathbf{n}}(u,u_b) = \begin{cases} \mathbf{f}(u_b)\cdot\mathbf{n} & \text{if } \nabla_u \mathbf{f}(u)\cdot\mathbf{n} > 0 \\ \mathbf{f}(u)\cdot\mathbf{n} & \text{else.} \end{cases}$$

In a first part, we present the class of schemes (nicknamed as Residual Distribution Schemes or RD or RDS for short) we are interested in, and show their link with more classical methods such as finite element ones. Then we recall a condition that guarantees that the numerical solution will converge to a weak solution of the problem. In the third part, we show that the RD schemes are also finite volume schemes: we compute explicitly the flux. In the fourth part, show that the now classical condition given by Tadmor in [17, 18] in one dimension fits very naturally in our framework.

2 Notations

From now on, we assume that Ω has a polyhedric boundary. This simplification is by no mean essential. We denote by $\mathcal{E}_h$ the set of internal edges/faces of $\mathcal{T}_h$, and by $\mathcal{F}_h$ those contained in $\partial\Omega$. $\mathcal{K}$ stands either for an element K or a face/edge $e \in \mathcal{E}_h \cup \mathcal{F}_h$. The boundary faces/edges are denoted by Γ. The mesh is assumed to be shape regular, h_K represents the diameter of the element K. Similarly, if $e \in \mathcal{E}_h \cup \mathcal{F}_h$, h_e represents its diameter.

Throughout this paper, we follow Ciarlet's definition [19, 20] of a finite element approximation: we have a set of degrees of freedom Σ_K of linear forms acting on the set $\mathbb{P}^k$ of polynomials of degree k such that the linear mapping

$$q \in \mathbb{P}^k \mapsto \big(\sigma_1(q),\ldots,\sigma_{|\Sigma_K|}(q)\big)$$

is one-to-one. The space $\mathbb{P}^k$ is spanned by the basis function $\{\varphi_\sigma\}_{\sigma\in\Sigma_K}$ defined by

$$\forall \sigma,\, \sigma', \sigma(\varphi_{\sigma'}) = \delta_\sigma^{\sigma'}.$$

We have in mind either Lagrange interpolations where the degrees of freedom are associated to points in K, or other type of polynomials approximation such as Bézier polynomials where we will also do the same geometrical identification. Considering all the elements covering Ω, the set of degrees of freedom is denoted by $\mathcal{S}$ and a generic degree of freedom by σ. We note that for any K,

$$\forall \mathbf{x} \in K, \quad \sum_{\sigma\in K} \varphi_\sigma(\mathbf{x}) = 1.$$

For any element K, $\#K$ is the number of degrees of freedom in K. If Γ is a face or a boundary element, $\#\Gamma$ is also the number of degrees of freedom in Γ.

The integer k is assumed to be the same for any element. We define

$$\mathcal{V}^h = \bigoplus_K \{v \in L^2(K), v_{|K} \in \mathbb{P}^k\}.$$

The solution will be sought for in a space V^h that is:

- Either $V^h = \mathcal{V}^h$. In that case, the elements of V^h can be discontinuous across internal faces/edges of $\mathcal{T}_h$. There is no conformity requirement on the mesh.
- Or $V^h = \mathcal{V}_h \cap C^0(\Omega)$ in which case the mesh needs to be conformal.

Throughout the text, we need to integrate functions. This is done via quadrature formula, and the symbol $\oint$ used in volume integrals

$$\oint_K v(\mathbf{x})\, d\mathbf{x}$$

or boundary integrals

$$\oint_{\partial K} v(\mathbf{x})\, d\gamma$$

means that these integrals are done via user defined numerical quadratures.

If $e \in \mathcal{E}_h$, represents any *internal* edge, i.e. $e \subset K \cap K^+$ for two elements K and K^+, we define for any function ψ the jump $[\nabla\psi] = \nabla\psi_{|K} - \nabla\psi_{|K^+}$. Here the choice of K and K^+ is important, hence also see relation (40) in section 5.2 where these element are defined in the relevant context. Similarly, $\{v\} = \frac{1}{2}\big(v_{|K} + v_{|K^+}\big)$.

If $\mathbf{x}$ and $\mathbf{y}$ are two vectors of $\mathbb{R}^q$, for q integer, $\langle \mathbf{x}, \mathbf{y}\rangle$ is their scalar product. In some occasions, it can also be denoted as $\mathbf{x}\cdot\mathbf{y}$ or $\mathbf{x}^T\mathbf{y}$. We also use $\mathbf{x}\cdot\mathbf{y}$ when $\mathbf{x}$ is a matrix and $\mathbf{y}$ a vector: it is simply the matrix-vector multiplication.

In sections 4 and 5, we have to deal with oriented graph. Given two vertices of this graph σ and σ', we write $\sigma > \sigma'$ to say that $[\sigma, \sigma']$ is a direct edge.

3 Schemes and conservation

3.1 Schemes

We begin this section by recalling the notion of flux. Let us consider any common edge or face Γ of K^+ and K^-, two elements. Let $\mathbf{n}$ be the normal to Γ, see Figure 1. Depending on the context, $\mathbf{n}$ is a scaled normal or $||\mathbf{n}|| = 1$. The symbols $S^\pm$ represent set of states, where S^+ is associated to K^+ and S^- to K^-. A flux

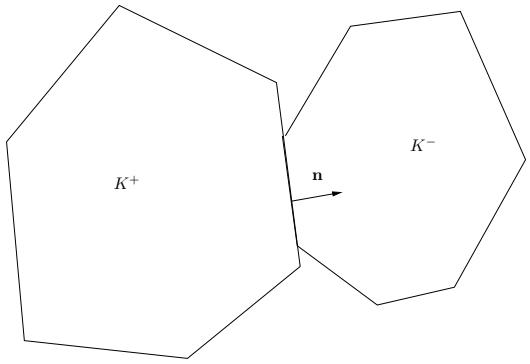

Figure 1: Geometrical setting

$\hat{\mathbf{f}}_{\mathbf{n}}(S^+, S^-)$ between K^+ and K^- has to satisfy

$$\hat{\mathbf{f}}_{\mathbf{n}}(S^+, S^-) = -\hat{\mathbf{f}}_{-\mathbf{n}}(S^-, S^+). \tag{3a}$$

and the consistency condition when the sets $S^\pm$ reduce to u

$$\hat{\mathbf{f}}_{\mathbf{n}}(S, S) = \mathbf{f}(u)\cdot\mathbf{n}. \tag{3b}$$

For a first order finite volume scheme, we have $S^+ = u_{K_+}$ and $S^- = u_{K_-}$, the average values of u in K^+ and K^-. For the other schemes, for example high order schemes, the definition is more involved.

In order to integrate the steady version of (1) on a domain $\Omega \subset \mathbb{R}^d$ with the boundary conditions (1b), on each element K and any degree of freedom $\sigma \in \mathcal{S}$ belonging to K, we define residuals $\Phi_\sigma^K(u^h)$. Following [11, 13], they are assumed to satisfy the following conservation relations: For any element K,

$$\sum_{\sigma\in K} \Phi_\sigma^K(u^h) = \int_{\partial K} \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma, \tag{4}$$

where $u^{h,-}$ is the approximation of the solution on the other side of the local edge/face of K. Note that in the case of a conformal mesh and with globally continuous elements, the condition reduces to

$$\sum_{\sigma\in K}\Phi_\sigma^K(u^h)=\int_{\partial K}\mathbf{f}(u^h)\cdot\mathbf{n}\,d\gamma.$$

Similarly, we consider residuals on the boundary elements Γ. On any such Γ, for any degree of freedom $\sigma\in\mathcal{S}\cap\Gamma$, we consider boundary residuals $\Phi_\sigma^\Gamma(u^h)$ that will satisfy the conservation relation

$$\sum_{\sigma\in\Gamma}\Phi_\sigma^\Gamma(u_h)=\int_\Gamma\left(\mathcal{F}_\mathbf{n}(u^h,u_b)-\mathbf{f}(u^h)\cdot\mathbf{n}\right)d\gamma. \tag{5}$$

Once this is done, the discretisation of (1) is achieved via: for any $\sigma\in\mathcal{S}$,

$$\sum_{K\subset\Omega,\sigma\in K}\Phi_\sigma^K(u^h)+\sum_{\Gamma\subset\partial\Omega,\sigma\in\Gamma}\Phi_\sigma^\Gamma(u^h)=0. \tag{6}$$

In (6), the first term represents the contribution of the internal elements. The second exists if $\sigma\in\partial\Omega$ and represents the contribution of the boundary conditions.

In fact, the formulation (6) is very natural. Consider a variational formulation of the steady version of (1):

$$\text{find } u^h\in V^h \text{ such that for any } v^h\in V^h, a(u^h,v^h)=0.$$

Let us show on three examples that this variational formulation leads to (6). They are

- The SUPG [21] variational formulation, with $u^h,v^h\in V^h=\mathcal{V}^h\cap C^0(\Omega)$:

$$\begin{aligned}a(u^h,v^h):=&-\int_\Omega\nabla v^h\cdot\mathbf{f}(u^h)\,d\mathbf{x}+\sum_{K\subset\Omega}h_K\int_K\left[\nabla\mathbf{f}(u^h)\cdot\nabla v^h\right]\tau_K\left[\nabla\mathbf{f}(u^h)\cdot\nabla u^h\right]d\mathbf{x}\\&+\int_{\partial\Omega}v^h\left(\mathcal{F}_\mathbf{n}(u^h,u_b)-\mathbf{f}(u^h)\cdot\mathbf{n}\right)d\gamma.\end{aligned} \tag{7}$$

  Here τ_K is a positive parameter.

- The Galerkin scheme with jump stabilization, see [22] for details. We have

$$\begin{aligned}a(u^h,v^h):=&-\int_\Omega\nabla v^h\cdot\mathbf{f}(u^h)\,d\mathbf{x}+\sum_{e\subset\Omega}\theta_e h_e^2\int_e\left[\nabla v^h\right]\cdot\left[\nabla u^h\right]d\gamma\\&+\int_{\partial\Omega}v^h\left(\mathcal{F}_\mathbf{n}(u^h,u_b)-\mathbf{f}(u^h)\cdot\mathbf{n}\right)d\gamma.\end{aligned} \tag{8}$$

  Here, $u^h,v^h\in V^h=\mathcal{V}^h\cap C^0(\Omega)$, and θ_e is a positive parameter.

- The discontinuous Galerkin formulation: we look for $u^h,v^h\in V^h=\mathcal{V}^h$ such that

$$a(u^h,v^h):=\sum_{K\subset\Omega}\left(-\int_K\nabla v^h\cdot\mathbf{f}(u^h)d\mathbf{x}+\int_{\partial K}v^h\cdot\hat{\mathbf{f}}_\mathbf{n}(u^h,u^{h,-})\,d\gamma\right). \tag{9}$$

  In (9), the boundary integral is a sum of integrals on the faces of K, and here for any face of K $u^{h,-}$ represents the approximation of u on the other side of that face in the case of internal elements, and u_b when that face is on $\partial\Omega$. Note that to fully comply with (6), we should have defined for boundary faces $u^{h,-}=u^h$, and then (9) is rewritten as

$$a(u^h,v^h):=\sum_{K\subset\Omega}\left(-\int_K\nabla v^h\cdot\mathbf{f}(u^h)d\mathbf{x}+\int_{\partial K}v^h\,\hat{\mathbf{f}}_\mathbf{n}(u^h,u^{h,-})\,d\gamma\right)+\sum_{\Gamma\subset\partial\Omega}\int_\Gamma v^h\cdot\left(\mathcal{F}_\mathbf{n}(u^h,u_b)-\mathbf{f}(u^h)\cdot\mathbf{n}\right)d\gamma. \tag{10}$$

  In (9), we have implicitly assumed $\hat{\mathbf{f}}_\mathbf{n}=\mathcal{F}_\mathbf{n}$ on the boundary edges.

In the SUPG, Galerkin scheme with jump stabilisation or the DG scheme, the boundary flux can be chosen different from $\mathcal{F}$. This can lead to boundary layers if these flux are not "enough" upwind, but we are not interested in these issues here.

Using the fact that the basis functions that span V_h have a *compact* support, then each scheme can be rewritten in the form (6) with the following expression for the residuals:

- For the SUPG scheme (7), the residual are defined by

$$\Phi_\sigma^K(u^h) = \int_{\partial K} \varphi_\sigma \mathbf{f}(u^h)\cdot\mathbf{n}\, d\gamma - \int_K \nabla\varphi_\sigma\cdot\mathbf{f}(u^h)\, d\mathbf{x} + h_K \int_K \Big(\nabla_u \mathbf{f}(u^h)\cdot\nabla\varphi_\sigma\Big)\tau_K\Big(\nabla_u\mathbf{f}(u^h)\cdot\nabla u^h\Big)\, d\mathbf{x}. \tag{11}$$

- For the Galerkin scheme with jump stabilization (8), the residuals are defined by:

$$\Phi_\sigma^K(u^h) = \int_{\partial K} \varphi_\sigma \mathbf{f}(u^h)\cdot\mathbf{n}\, d\gamma - \int_K \nabla\varphi_\sigma\cdot\mathbf{f}(u^h)\, d\mathbf{x} + \sum_{e \text{ faces of } K} \frac{\theta_e}{2} h_e^2 \int_{\partial K} [\nabla u^h]\cdot[\nabla\varphi_\sigma]\, d\gamma \tag{12}$$

  with $\theta_e > 0$. Here, since the mesh is conformal, any internal edge e (or face in 3D) is the intersection of the element K and another element denoted by K^+.

- For the discontinuous Galerkin scheme,

$$\Phi_\sigma^K(u^h) = -\int_K \nabla\varphi_\sigma\cdot\mathbf{f}(u^h) d\mathbf{x} + \int_{\partial K} \varphi_\sigma\cdot\hat{\mathbf{f}}_\mathbf{n}(u^h, u^{h,-})\, d\gamma \tag{13}$$

  using the second definition of $u^{h,-}$.

- The boundary residuals are

$$\Phi_\sigma^\Gamma(u^h) = \int_\Gamma \varphi_\sigma\big(\mathcal{F}_\mathbf{n}(u^h, u_b) - \mathbf{f}(u^h)\cdot\mathbf{n}\big)\, d\gamma \tag{14}$$

All these residuals satisfy the relevant conservation relations, namely (4) or (5), depending if we are dealing with element residuals or boundary residuals.

For now, we are just rephrasing classical finite element schemes into a purely numerical framework. However, considering the pure numerical point of view and forgetting the variational framework, we can go further and define schemes that have no clear variational formulation. These are the limited Residual Distributive Schemes, see [11, 13], namely

$$\Phi_\sigma^K(u^h) = \beta_\sigma \int_{\partial K} \mathbf{f}(u^h)\cdot\mathbf{n}\, d\gamma \tag{15}$$

or

$$\Phi_\sigma^K(u^h) = \beta_\sigma \int_{\partial K} \mathbf{f}(u^h)\cdot\mathbf{n}\, d\gamma + \theta_K h_K \int_K \Big(\nabla_u \mathbf{f}(u^h)\cdot\nabla\varphi_\sigma\Big)\tau_K\Big(\nabla_u\mathbf{f}(u^h)\cdot\nabla u^h\Big)\, d\mathbf{x}, \qquad \theta_K \geq 0 \tag{16}$$

or

$$\Phi_\sigma^K(u^h) = \beta_\sigma \int_{\partial K} \mathbf{f}(u^h)\cdot\mathbf{n}\, d\gamma + \theta_e\, h_e^2 \int_{\partial K} [\nabla u^h]\cdot[\nabla\varphi_\sigma]\, d\gamma \qquad \theta_e \geq 0 \tag{17}$$

where the parameters β_σ are defined to guarantee conservation,

$$\sum_{\sigma\in K} \beta_\sigma = 1$$

and such that (16) without the streamline term and (17) without the jump term satisfy a discrete maximum principle. The streamline term and jump term are introduced because one can easily see that spurious modes may exist, but their role is very different compared to (11) and (12) where they are introduced to stabilize the Galerkin scheme: if formally the maximum principle is violated, experimentally the violation is extremely small if existent at all. See [7, 11] for more details.

A similar construction can be done starting from a discontinuous Galerkin scheme, see [10, 9]. A second order version is described in appendix A.

The non-linear stability is provided by the coefficient β_σ which is a non-linear function of u^h. Possible values of β_σ are described in remark 3.1 bellow.

Remark 3.1. *The coefficients β_σ introduced in the relations (16) and (17) are defined by:*

$$\beta_\sigma = \frac{\max(0, \frac{\Phi_\sigma}{\Phi})}{\sum_{\sigma' \in K} \max(0, \frac{\Phi_{\sigma'}}{\Phi})}. \tag{18}$$

These coefficients are always defined and garantee a local maximum principle for (16) and (17): this is again a consequence of the conservation properties, see e.g. [11]. Note this is true for any order of interpolation.

3.2 Conservation

From (6), using the conservation relations (5) and (4), we obtain for any $v^h \in V^h$,

$$v_h = \sum_{\sigma \in \mathcal{S}} v_\sigma \varphi_\sigma,$$

the following relation:

$$\begin{aligned} 0 = &- \int_\Omega \nabla v_h \cdot \mathbf{f}(u^h)\, d\mathbf{x} + \int_{\partial\Omega} v^h \big(\hat{\mathbf{f}}_\mathbf{n}(u^h, u_b) - \mathbf{f}(u^h)\cdot\mathbf{n}\big)\, d\gamma \\ &+ \sum_{e \in \mathcal{E}_h} \int_e [v^h] \hat{\mathbf{f}}_\mathbf{n}(u^h, u^{h,-})\, d\gamma + \sum_{K\subset\Omega} \frac{1}{\#K} \Big(\sum_{\sigma,\sigma' \in K} (v_\sigma - v_{\sigma'}) \Big(\Phi_\sigma^K(u^h) - \Phi_\sigma^{K,Gal}(u^h) \Big) \Big) \\ &+ \sum_{\Gamma \subset \partial\Omega} \frac{1}{\#\Gamma} \Big(\sum_{\sigma,\sigma' \in \Gamma} (v_\sigma - v_{\sigma'}) (\Phi_\sigma^\Gamma(u^h, u_b) - \Phi_\sigma^{Gal,\Gamma}(u^h, u_b)) \Big) \end{aligned} \tag{19}$$

where

$$\Phi_\sigma^{K,Gal}(u^h) = -\int_K \nabla\varphi_\sigma \cdot \mathbf{f}(u^h)\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma \hat{\mathbf{f}}_\mathbf{n}(u^h, u^{h,-})\, d\gamma, \qquad \Phi_\sigma^{\Gamma,Gal}(u^h, u_b) = \int_\Gamma \varphi_\sigma \big(\hat{\mathbf{f}}_\mathbf{n}(u^h, u_b) - \mathbf{f}(u^h)\cdot\mathbf{n}\big)\, d\gamma.$$

Proof. We start from (6) which is multiplied by v_σ, and these relations are added for each $\sigma \in \mathcal{S}$. We get:

$$0 = \sum_{\sigma \in \mathcal{S}} v_\sigma \Big(\sum_{K\subset\Omega, \sigma\in K} \Phi_\sigma^K(u^h) + \sum_{\Gamma\subset\partial\Omega, \sigma\in\Gamma} \Phi_\sigma^\Gamma(u^h, u_b) \Big).$$

Permuting the sums on σ and K, then on σ and Γ, we get:

$$0 = \sum_{K\subset\Omega} \Big(\sum_{\sigma\in K} v_\sigma \Phi_\sigma^K(u^h) \Big) + \sum_{\Gamma\subset\partial\Omega} \Big(\sum_{\sigma\in\Gamma} v_\sigma \Phi_\sigma^\Gamma(u^h, u_b) \Big).$$

We look at the first term, the second is done similarly. We have, introducing $\Phi_\sigma^{K,Gal}$ and $\#K$ the number of degrees of freedom in K,

$$\begin{aligned}\sum_{\sigma\in K} v_\sigma \Phi_\sigma^K(u^h) &= \sum_{\sigma\in K} v_\sigma \Phi_\sigma^{K,Gal}(u^h) + \sum_{\sigma\in K} v_\sigma \Big(\Phi_\sigma^K(u^h) - \Phi_\sigma^{K,Gal}(u^h)\Big)\\ &= -\int_K \nabla v_h \cdot \mathbf{f}(u^h)\, d\mathbf{x} + \int_{\partial K} v^h \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma + \sum_{\sigma\in K} v_\sigma \Big(\Phi_\sigma^K(u^h) - \Phi_\sigma^{K,Gal}(u^h)\Big)\\ &= -\int_K \nabla v_h \cdot \mathbf{f}(u^h)\, d\mathbf{x} + \int_{\partial K} v^h \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma + \frac{1}{\#K}\sum_{\sigma,\sigma'\in K} (v_\sigma - v_{\sigma'})\Big(\Phi_\sigma^K(u^h) - \Phi_\sigma^{K,Gal}(u^h)\Big)\end{aligned}$$

because

$$\sum_{\sigma\in K} \big(\Phi_\sigma^K(u^h) - \Phi_\sigma^{K,Gal}(u^h)\big) = 0.$$

Similarly, we have

$$\sum_{\sigma\in\Gamma} v_\sigma \Phi_\sigma^\Gamma(u^h) = \int_\Gamma v^h \big(\hat{\mathbf{f}}_{\mathbf{n}}(u^h, u_b) - \mathbf{f}(u^h)\cdot\mathbf{n}\big)\, d\gamma + \sum \frac{1}{\#\Gamma}\sum_{\sigma,\sigma'\in\Gamma} (v_\sigma - v_{\sigma'})(\Phi_\sigma^\Gamma(u^h, u_b) - \Phi_\sigma^{Gal,\Gamma}(u^h, u_b))$$

Adding all the relations, we get:

$$\begin{aligned}0 = &\sum_{K\subset\Omega}\Big(-\int_K \nabla v_h\cdot\mathbf{f}(u^h)\, d\mathbf{x} + \int_{\partial K} v^h \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma\Big) + \sum_{\Gamma\subset\partial\Omega}\int_\Gamma v^h\big(\hat{\mathbf{f}}_{\mathbf{n}}(u^h, u_b) - \mathbf{f}(u^h)\cdot\mathbf{n}\big)\, d\gamma\\ &+ \sum_{K\subset\Omega}\frac{1}{\#K}\Big(\sum_{\sigma,\sigma'\in K}(v_\sigma - v_{\sigma'})\Big(\Phi_\sigma^K(u^h) - \Phi_\sigma^{K,Gal}(u^h)\Big)\Big)\\ &+ \sum_{\Gamma\subset\partial\Omega}\frac{1}{\#\Gamma}\Big(\sum_{\sigma,\sigma'\in\Gamma}(v_\sigma - v_{\sigma'})(\Phi_\sigma^\Gamma(u^h, u_b) - \Phi_\sigma^{Gal,\Gamma}(u^h, u_b))\Big)\end{aligned}$$

i.e. after having defined $[v^h] = v^h - v^{h,-}$ and chosen one orientation of the internal edges $e \in \mathcal{E}_h$, we get (19). □

The relation (19) is instrumental in proving the following results. The first one is proved in [8], and is a generalisation of the classical Lax-Wendroff theorem.

Theorem 3.2. *Assume the family of meshes $\mathcal{T} = (\mathcal{T}_h)$ is shape regular. We assume that the residuals $\{\Phi_\sigma^{\mathcal{K}}\}_{\sigma\in\mathcal{K}}$, for $\mathcal{K}$ an element or a boundary element of $\mathcal{T}_h$, satisfy:*

- *For any $M \in \mathbb{R}^+$, there exists a constant C which depends only on the family of meshes $\mathcal{T}_h$ and M such that for any $u^h \in V^h$ with $||u^h||_\infty \le M$, then*

$$\big|\Phi_\sigma^{\mathcal{K}}(u^h{}_{|\mathcal{K}})\big| \le C \sum_{\sigma,\sigma'\in\mathcal{K}} |u_\sigma^h - u_{\sigma'}^h|$$

- *The conservation relations (4) and (5).*

Then if there exists a constant C_{max} such that the solutions of the scheme (6) satisfy $||u^h||_\infty \le C_{max}$ and a function $v \in L^2(\Omega)$ such that $(u^h)_h$ or at least a sub-sequence converges to v in $L^2(\Omega)$, then v is a weak solution of (1)

Proof. The proof can be found in [8], it uses (19) and some adaptation of the ideas of [23]. One of the key arguments comes from the consistency of the flux $\hat{\mathbf{f}}$ as well as (3a) □

Another consequence of (19) is the following result on entropy inequalities:

Proposition 3.3. *Let $(U, \mathbf{g})$ be a entropy-flux couple for (1) and $\hat{\mathbf{g}}_\mathbf{n}$ be a numerical entropy flux consistent with $\mathbf{g}\cdot\mathbf{n}$. Assume that the residuals satisfy: for any element K,*

$$\sum_{\sigma\in K}\langle \nabla_u U(u_\sigma), \Phi_\sigma^K\rangle \geq \int_{\partial K} \hat{\mathbf{g}}_\mathbf{n}(u^h, u^{h,-})\, d\gamma \tag{20a}$$

and for any boundary edge e,

$$\sum_{\sigma\in e}\langle \nabla_u U(u_\sigma), \Phi_\sigma^e\rangle \geq \int_e \left(\hat{\mathbf{g}}_\mathbf{n}(u^h, u_b) - \mathbf{g}(u^h)\cdot\mathbf{n}\right) d\gamma. \tag{20b}$$

Then, under the assumptions of theorem 3.2, the limit weak solution also satisfies the following entropy inequality: for any $\varphi \in C^1(\overline{\Omega})$, $\varphi \geq 0$,

$$-\int_\Omega \nabla\varphi\cdot\mathbf{g}(u)\, d\mathbf{x} + \int_{\partial\Omega^-} \varphi\, \mathbf{g}(u_b)\cdot\mathbf{n}\, d\gamma \leq 0.$$

Proof. The proof is similar to that of theorem 3.2. □

Another consequence of (19) is the following condition under which one can guarantee to have a $k+1$-th order accurate scheme. We first introduce the (weak) truncation error

$$\mathcal{E}(u^h, \varphi) = \sum_{\sigma\in\mathcal{S}_h} \varphi_\sigma \left[\sum_{K\subset\Omega, \sigma\in K} \Phi_\sigma^K + \sum_{\Gamma\subset\partial\Omega, \sigma\in\Gamma} \Phi_\sigma^\Gamma\right]. \tag{21}$$

If the solution of the *steady problem* u is smooth enough and the residuals, computed with the interpolant $\pi_h(u)$ of the solution, are such that for any element K and boundary element Γ

$$\Phi_\sigma^K(\pi_h(u)) = \mathcal{O}\big(h^{k+d}\big), \qquad \Phi_\sigma^\Gamma(\pi_h(u)) = \mathcal{O}\big(h^{k+d-1}\big) \tag{22}$$

and if the approximation $\mathbf{f}\big(u^h\big)$ of $\mathbf{f}(u)$ is accurate with order $k+1$, then the truncation error satisfies the following relation

$$|\mathcal{E}\big(\pi_h(u), \varphi\big)| \leq C(\mathbf{f}, u)\ ||\varphi||_{H^1(\Omega)}\ h^{k+1},$$

with C a constant which depends only on $\mathbf{f}$, and $||u||_\infty$.

Proof. We first show that $\Phi_\sigma^{K,Gal}(\pi_h(u)) = \mathcal{O}\big(h^{k+d}\big)$. Since u is regular enough, we have pointwise $\operatorname{div}\mathbf{f}(u) = 0$ on K, so that, by consistency of the flux,

$$0 = -\int_K \nabla\varphi\cdot\mathbf{f}(u)\, d\mathbf{x} + \int_{\partial K} \varphi\hat{\mathbf{f}}_\mathbf{n}(u,u)\, d\gamma.$$

Then,

$$\begin{aligned}
\Phi_\sigma^{K,Gal}(\pi_h(u)) &= -\int_K \nabla\varphi_\sigma\cdot\big(\mathbf{f}(\pi_h(u)) - \mathbf{f}(u)\big)\, d\mathbf{x} + \int_{\partial K}\varphi_\sigma\big(\hat{\mathbf{f}}_\mathbf{n}(\pi_h(u),\pi_h(u)) - \hat{\mathbf{f}}_\mathbf{n}(u,u)\big)\, d\gamma \\
&= |K|\times\mathcal{O}(h^{-1})\times\mathcal{O}(h^{k+1}) + |\partial K|\times\mathcal{O}(1)\mathcal{O}(h^{k+1}) \\
&= \mathcal{O}(h^d)\times\mathcal{O}(h^{-1})\times\mathcal{O}(h^{k+1}) + \mathcal{O}(h^{d-1})\times\mathcal{O}(1)\times\mathcal{O}(h^{k+1}) \\
&= \mathcal{O}(h^{d+k})
\end{aligned}$$

because the flux is Lipschitz continuous and the mesh is regular.

The result on the boundary term is similar since the boundary numerical flux is upwind and the boundary of Ω is not characteristic: only two types of boundary faces exists, the upwind and downwind ones. On the downwind faces, the boundary flux vanishes. On the upwind ones, we get the estimate for the Galerkin boundary residuals thanks to the same approximation argument.

The mesh is assumed to be regular: the number of elements (resp. edges) is $O(h^{-d})$ (resp. $O(h^{-d+1})$). Let us assume (22). Let $v \in C_0^1(\overline{\Omega})$. Using (19) for $\pi_h(u)$,

$$\begin{aligned}\mathcal{E}\big(u^h,\varphi\big) = &-\int_\Omega \nabla\pi_h(v)\cdot\mathbf{f}(\pi_h(u))\,d\mathbf{x} + \int_{\partial\Omega}\pi_h(v)\big(\hat{\mathbf{f}}_{\mathbf{n}}(\pi_h(u),u_b)-\mathbf{f}(\pi_h(u))\cdot\mathbf{n}\big)\,d\gamma\\ &+\sum_{e\in\mathcal{E}_h}\int_e[\pi_h(v)]\hat{\mathbf{f}}_{\mathbf{n}}(\pi_h(u),\pi_h(u)^-)\,d\gamma+\sum_{K\subset\Omega}\frac{1}{\#K}\Big(\sum_{\sigma,\sigma'\in K}(v_\sigma-v_{\sigma'})\Big(\Phi_\sigma^K(\pi_h(u))-\Phi_\sigma^{K,Gal}(\pi_h(u))\Big)\Big)\\ &+\sum_{\Gamma\subset\partial\Omega}\frac{1}{\#\Gamma}\Big(\sum_{\sigma,\sigma'\in\Gamma}(v_\sigma-v_{\sigma'})(\Phi_\sigma^\Gamma\big(\pi_h(u),u_b\big)-\Phi_\sigma^{Gal,\Gamma}(\pi_h(u),u_b))\Big).\end{aligned}$$

where $\pi_h(u)^-$ represents the interpolant of u on K^-.

We have, using

$$-\int_\Omega\nabla\pi_h(v)\cdot\mathbf{f}(u)\,d\mathbf{x}+\int_{\partial\Omega}\pi_h(v)\big(\mathcal{F}_{\mathbf{n}}(u,u_b)-\mathbf{f}(u)\cdot\mathbf{n}\big)\,d\gamma=0,$$

$$\begin{aligned}&-\int_\Omega\nabla\pi_h(v)\cdot\mathbf{f}(\pi_h(u))\,d\mathbf{x}+\int_{\partial\Omega}\pi_h(v)\big(\hat{\mathbf{f}}_{\mathbf{n}}(\pi_h(u),u_b)-\mathbf{f}(\pi_h(u))\cdot\mathbf{n}\big)\,d\gamma\\ &\quad=-\int_\Omega\nabla\pi_h(v)\cdot\big(\mathbf{f}(\pi_h(u))-\mathbf{f}(u)\big)\,d\mathbf{x}+\int_{\partial\Omega}\pi_h(v)\big(\hat{\mathbf{f}}_{\mathbf{n}}(\pi_h(u),u_b)-\hat{\mathbf{f}}_{\mathbf{n}}(u,u_b)\big)\,d\gamma\\ &\qquad\quad-\int_{\partial\Omega}\pi_h(v)\big(\mathbf{f}(\pi_h(u))-\mathbf{f}(u)\big)\cdot\mathbf{n}\,d\gamma\\ &\quad=\mathcal{O}(h^{k+1})\end{aligned}$$

since the flux on the boundary is the upwind flux $\mathcal{F}_{\mathbf{n}}$, and using the approximation properties of $\pi_h(u)$.

Then

$$\sum_{e\in\mathcal{E}_h}\int_e[\pi_h(v)]\hat{\mathbf{f}}_{\mathbf{n}}(\pi_h(u),\pi_h(u)^-)\,d\gamma=\mathcal{O}(h^{-d+1})\times\mathcal{O}(h^{d-1})\times\mathcal{O}(h^{k+1})\times\mathcal{O}(1)=\mathcal{O}(h^{k+1}),$$

$$\sum_{K\subset\Omega}\frac{1}{\#K}\Big(\sum_{\sigma,\sigma'\in K}(v_\sigma-v_{\sigma'})\Big(\Phi_\sigma^K(\pi_h(u))-\Phi_\sigma^{K,Gal}(\pi_h(u))\Big)\Big)=\mathcal{O}(h^{-d})\times\mathcal{O}(h)\times\mathcal{O}(h^{k+d})=\mathcal{O}(h^{k+1}),$$

and similarly

$$\sum_{\Gamma\subset\partial\Omega}\frac{1}{\#\Gamma}\Big(\sum_{\sigma,\sigma'\in\Gamma}(v_\sigma-v_{\sigma'})(\Phi_\sigma^\Gamma\big(\pi_h(u),u_b\big)-\Phi_\sigma^{Gal,\Gamma}(\pi_h(u),u_b))\Big)=\mathcal{O}(h^{-d+1})\times\mathcal{O}(h)\times\mathcal{O}(h^{k+d-1})=\mathcal{O}(h^{k+1})$$

thanks to the regularity of the mesh, that $\pi_h(v)$ is the interpolant of a C^1 function and the previous estimates. □

Remark 3.4 (Numerical integration)**.** *In practice, the integrals are evaluated by numerical integration. The results still holds true provided the quadrature formula are of order $k+1$. This is in contrast with the common practice, but let us emphasis this is valid only for steady problems. However, similar arguments can be developed for unsteady problems, see [12, 15].*

4 Flux formulation of Residual Distribution schemes

In this section we show that the scheme (6) also admits a flux formulation, with an explicit form of the flux: the method is also locally conservative. Local conservation is of course well known for the Finite Volume and discontinuous Galerkin approximations. It is much less understood for the continuous finite elements methods, despite the papers [21, 24]. Referring to (3), the aim of this section is to define $\hat{\mathbf{f}}$ and $S^{\pm}$ in the RDS case.

We first show why a finite volume can be reinterpreted as an RD scheme. This helps to understand the structure of the problem. Then we show that any RD scheme can be equivalently rephrased as a finite volume scheme, we explicitly provide the flux formula as well as the control volumes. In order to illustrate this result, we give several examples: the general RD scheme with $\mathbb{P}^1$ and $\mathbb{P}^2$ approximation on simplex, the case of a $\mathbb{P}^1$ RD scheme using a particular form of the residuals so that one can better see the connection with more standard formulations, and finally an example with a discontinuous Galerkin formulation using $\mathbb{P}^1$ approximation.

4.1 Finite volume as Residual distribution schemes

Here, we rephrase [6]. The notations are defined in Figure 2. Again, we specialize ourselves to the case

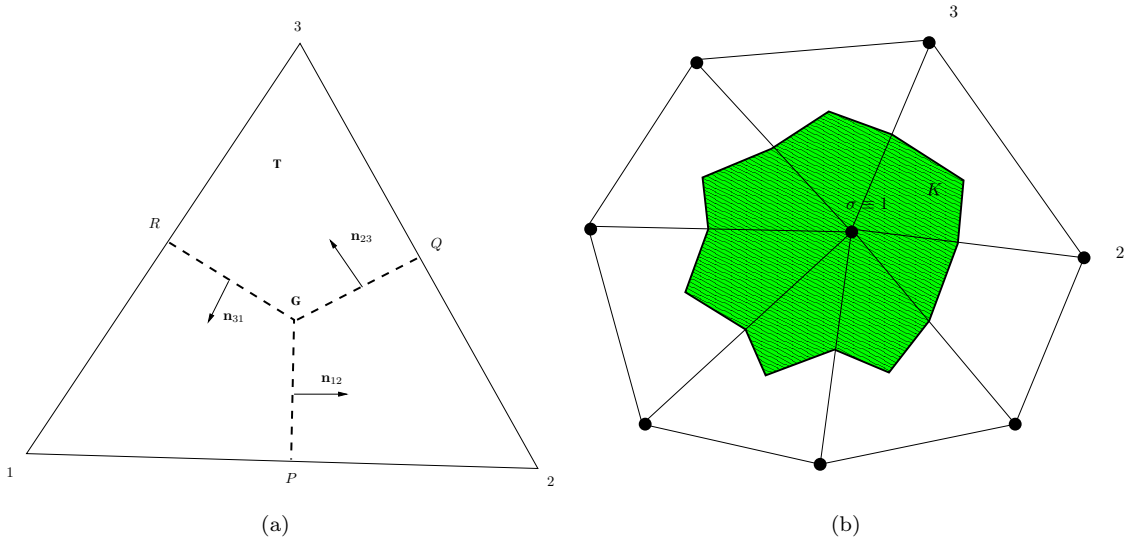

Figure 2: Notations for the finite volume schemes. On the left: definition of the control volume for the degree of freedom σ. The vertex σ plays the role of the vertex 1 on the left picture for the triangle K. The control volume C_σ associated to $\sigma = 1$ is green on the right and corresponds to $1PGR$ on the left. The vectors $\mathbf{n}_{ij}$ are normal to the internal edges scaled by the corresponding edge length

of triangular elements, but *exactly the same arguments* can be given for more general elements, provided a conformal approximation space can be constructed. This is the case for triangle elements, and we can take $k = 1$.

The control volumes in this case are defined as the median cell, see figure 2. We concentrate on the approximation of div $\mathbf{f}$, see equation (1). Since the boundary of C_σ is a closed polygon, the scaled outward normals $\mathbf{n}_\gamma$ to ∂C_σ sum up to 0:

$$\sum_{\gamma \subset \partial C_\sigma} \mathbf{n}_\gamma = 0$$

where γ is any of the segment included in ∂C_σ, such as PG on Figure 2. Hence

$$\begin{aligned}\sum_{\gamma\subset\partial C_\sigma}\hat{\mathbf{f}}_{\mathbf{n}_\gamma}(u_\sigma,u^-) &= \sum_{\gamma\subset\partial C_\sigma}\hat{\mathbf{f}}_{\mathbf{n}_\gamma}(u_\sigma,u^-) - \left(\sum_{\gamma\subset\partial C_\sigma}\mathbf{n}_\gamma\right)\cdot\mathbf{f}(u_\sigma)\\ &= \sum_{K,\sigma\in K}\sum_{\gamma\subset\partial C_\sigma\cap K}\left(\hat{\mathbf{f}}_{\mathbf{n}_\gamma}(u_\sigma,u^-)-\mathbf{f}(u_\sigma)\cdot\mathbf{n}_\gamma\right)\end{aligned}$$

To make things explicit, in K, the internal boundaries are PG, QG and RG, and those around $\sigma\equiv 1$ are PG and RG. We set

$$\begin{aligned}\Phi_\sigma^K(u^h) &= \sum_{\gamma\subset\partial C_\sigma\cap K}\left(\hat{\mathbf{f}}_{\mathbf{n}_\gamma}(u_\sigma,u^-)-\mathbf{f}(u_\sigma)\cdot\mathbf{n}_\gamma\right)\\ &= \sum_{\gamma\subset\partial(C_\sigma\cap K)}\hat{\mathbf{f}}_{\mathbf{n}_\gamma}(u_\sigma,u^-).\end{aligned}\tag{23}$$

The last relation uses the consistency of the flux and the fact that $C_\sigma\cap K$ is a closed polygon. The quantity $\Phi_\sigma^K(u^h)$ is the normal flux on $C_\sigma\cap K$. If now we sum up these three quantities and get:

$$\begin{aligned}\sum_{\sigma\in K}\Phi_\sigma^K(u_h) &= \left(\hat{\mathbf{f}}_{\mathbf{n}_{12}}(u_1,u_2)-\hat{\mathbf{f}}_{\mathbf{n}_{13}}(u_1,u_3)-\mathbf{f}(u_1)\cdot\mathbf{n}_{12}+\mathbf{f}(u_1)\cdot\mathbf{n}_{31}\right)\\ &+\left(\hat{\mathbf{f}}_{\mathbf{n}_{23}}(u_2,u_3)-\hat{\mathbf{f}}_{\mathbf{n}_{12}}(u_2,u_1)+\mathbf{f}(u_2)\cdot\mathbf{n}_{12}-\mathbf{f}(u_2)\cdot\mathbf{n}_{23}\right)\\ &+\left(-\hat{\mathbf{f}}_{\mathbf{n}_{23}}(u_3,u_2)+\hat{\mathbf{f}}_{\mathbf{n}_{31}}(u_3,u_1)-\mathbf{f}(u_3)\cdot\mathbf{n}_{23}+\mathbf{f}(u_3)\cdot\mathbf{n}_{31}\right)\\ &= \mathbf{f}(u_1)\cdot\left(\mathbf{n}_{12}-\mathbf{n}_{31}\right)+\mathbf{f}(u_2)\cdot\left(-\mathbf{n}_{23}+\mathbf{n}_{31}\right)+\mathbf{f}(u_3)\cdot\left(\mathbf{n}_{31}-\mathbf{n}_{23}\right)\\ &= \mathbf{f}(u_1)\cdot\frac{\mathbf{n}_1}{2}+\mathbf{f}(u_2)\cdot\frac{\mathbf{n}_2}{2}+\mathbf{f}(u_3)\cdot\frac{\mathbf{n}_3}{2}\end{aligned}$$

where $\mathbf{n}_j$ is the scaled inward normal of the edge opposite to vertex σ_j, i.e. twice the gradient of the $\mathbb{P}^1$ basis function φ_{σ_j} associated to this degree of freedom. Thus, we can reinterpret the sum as the boundary integral of the Lagrange interpolant of the flux. The finite volume scheme is then a residual distribution scheme with residual defined by (23) and a total residual defined by

$$\Phi^K := \int_{\partial K}\mathbf{f}^h\cdot\mathbf{n},\qquad \mathbf{f}^h=\sum_{\sigma\in K}\mathbf{f}(u_\sigma)\varphi_\sigma.\tag{24}$$

4.2 Residual distribution schemes as finite volume schemes.

In this section, we show how to interpret RD schemes as finite volume schemes. This amounts to defining control volumes and flux functions. We first have to adapt the notion of consistency. As recalled in the section 3.2, two of the key arguments in the proof of the Lax-Wendroff theorem are related to the structure of the flux, for classical finite volume schemes. In [8], the proof is adapted to the case of Residual Distribution schemes. The property that stands for the consistency is that if all the states are identical in an element, then each of the residuals vanishes. Hence, we define a multidimensional flux as follows:

Definition 4.1. *A multidimensional flux*

$$\hat{\mathbf{f}}_{\mathbf{n}} := \hat{\mathbf{f}}_{\mathbf{n}}(u_1,\ldots,u_N)$$

is consistent if, when $u_1=u_2=\ldots=u_N=u$ *then*

$$\hat{\mathbf{f}}_{\mathbf{n}}(u,\ldots,u)=\mathbf{f}(u)\cdot\mathbf{n}.$$

We proceed first with the general case and show the connection with elementary fact about graphs, and then provide several examples. The results of this section apply to any finite element method but also to discontinuous Galerkin methods. There is no need for exact evaluation of integral formula (surface or boundary), so that these results apply to schemes as they are implemented.

4.2.1 General case

One can deal with the general case, i.e when K is a polytope contained in $\mathbb{R}^d$ with degrees of freedoms on the boundary of K. The set $\mathcal{S}$ is the set of degrees of freedom. We consider a triangulation $\mathcal{T}_K$ of K whose vertices are exactly the elements of $\mathcal{S}$. Choosing an orientation of K, it is propagated on $\mathcal{T}_K$: the edges are oriented.

The problem is to find quantities $\hat{\mathbf{f}}_{\sigma,\sigma'}$ for any edge $[\sigma,\sigma']$ of $\mathcal{T}_K$ such that:

$$\Phi_\sigma = \sum_{\text{edges } [\sigma,\sigma']} \hat{\mathbf{f}}_{\sigma,\sigma'} + \hat{\mathbf{f}}^b_\sigma \tag{25a}$$

with

$$\hat{\mathbf{f}}_{\sigma,\sigma'} = -\hat{\mathbf{f}}_{\sigma',\sigma} \tag{25b}$$

and $\hat{\mathbf{f}}^b_\sigma$ is the 'part' of $\oint_{\partial K} \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma$ associated to σ. The control volumes will be defined by their normals so that we get consistency.

Note that (25b) implies the conservation relation

$$\sum_{\sigma\in K} \Phi_\sigma = \sum_{\sigma\in K} \hat{\mathbf{f}}^b_\sigma. \tag{25c}$$

In short, we will consider

$$\hat{\mathbf{f}}^b_\sigma = \oint_{\partial K} \varphi_\sigma\, \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma, \tag{25d}$$

but other examples can be considered provided the consistency (25c) relation holds true, see for example section 4.2.2. Any edge $[\sigma,\sigma']$ is either direct or, if not, $[\sigma',\sigma]$ is direct. Because of (25b), we only need to know $\hat{\mathbf{f}}_{\sigma,\sigma'}$ for direct edges. Thus we introduce the notation $\hat{\mathbf{f}}_{\{\sigma,\sigma'\}}$ for the flux assigned to the direct edge whose extremities are σ and σ'. We can rewrite (25a) as, for any $\sigma \in \mathcal{S}$,

$$\sum_{\sigma'\in\mathcal{S}} \varepsilon_{\sigma,\sigma'} \hat{\mathbf{f}}_{\{\sigma,\sigma'\}} = \Psi_\sigma := \Phi_\sigma - \hat{\mathbf{f}}^b_\sigma, \tag{26}$$

with

$$\varepsilon_{\sigma,\sigma'} = \begin{cases} 0 & \text{if } \sigma \text{ and } \sigma' \text{ are not on the same edge of } \mathcal{T}, \\ 1 & \text{if } [\sigma,\sigma'] \text{ is an edge and } \sigma \to \sigma' \text{ is direct}, \\ -1 & \text{if } [\sigma,\sigma'] \text{ is an edge and } \sigma' \to \sigma \text{ is direct}. \end{cases}$$

$\mathcal{E}^+$ represents the set of direct edges.

Hence the problem is to find a vector $\hat{\mathbf{f}} = (\hat{\mathbf{f}}_{\{\sigma,\sigma'\}})_{\{\sigma,\sigma'\} \text{ direct edges}}$ such that

$$A\hat{\mathbf{f}} = \Psi$$

where $\Psi = (\Psi_\sigma)_{\sigma\in\mathcal{S}}$ and $A_{\sigma\sigma'} = \varepsilon_{\sigma,\sigma'}$.

We have the following lemma which shows the existence of a solution.

Lemma 4.2. *For any couple $\{\Phi_\sigma\}_{\sigma\in\mathcal{S}}$ and $\{\hat{\mathbf{f}}^b_\sigma\}_{\sigma\in\mathcal{S}}$ satisfying the condition (25c), there exists numerical flux functions $\hat{\mathbf{f}}_{\sigma,\sigma'}$ that satisfy (25). Recalling that the matrix of the Laplacian of the graph is $L = AA^T$, we have*

1. *The rank of L is $|\mathcal{S}|-1$ and its image is $\big(span\{\mathbf{1}\}\big)^\perp$. We still denote the inverse of L on $\big(span\{\mathbf{1}\}\big)^\perp$ by L^{-1},*

2. *With the previous notations, a solution is*

$$\big(\hat{\mathbf{f}}_{\{\sigma,\sigma'\}}\big)_{\{\sigma,\sigma'\}\ direct\ edges} = A^T L^{-1}\big(\Psi_\sigma\big)_{\sigma\in\mathcal{S}}. \tag{27}$$

Proof. We first have $\mathbf{1}^T A = 0$: $\text{Im}\, A \subset \big(\text{span}\ \{1\}\big)^\perp (\subset \mathbb{R}^{|\mathcal{S}|})$. Let us show that we have equality. In order to show this, we notice that the matrix A is nothing more that the incidence matrix of the oriented graph $\mathcal{G}$ defined by the triangulation $\mathcal{T}$. It is known [25] that its null space of L is equal to the number of connected components of the graph, i.e. here $\dim\ker L = 1$. Since

$$L\,\mathbf{1} = 0,$$

we see that $\ker L = \text{span}\ \{\mathbf{1}\}$, so that $\text{Im}\, L = \big(\text{span}\ \{\mathbf{1}\}\big)^T$ because L is symmetric. We can define the inverse of L on $\text{Im}\, L$, denoted by L^{-1}.

Let $x \in \big(\text{span}\ \{\mathbf{1}\}\big)^\perp = \text{Im}\, L$. There exists $y \in \mathbb{R}^{|S|}$ such that $x = Ly = A(A^T y)$: this shows that $x \in \text{Im}\, A$ and thus $\text{Im}\, A = \big(\text{span}\ \{\mathbf{1}\}\big)^\perp = \big(\text{Im}\, L\big)^\perp$. From this we deduce that $\text{rank}\, A = |\mathcal{S}| - 1$ because $\text{Im}\, A \subset \mathbb{R}^{|\mathcal{S}|}$.

Let $\Psi \in \mathbb{R}^{|\mathcal{S}|}$ be such that $\langle \mathbf{1}, \Psi\rangle = 0$. We know there exists a unique $z \in \big(\text{span}\ \{\mathbf{1}\}\big)^\perp$ such that $Lz = \Psi$, i.e.

$$A(A^T z) = \Psi.$$

This shows that a solution is given by (27). □

This set of flux are consistent and we can estimate the normals $\mathbf{n}_{\sigma,\sigma'}$. In the case of a constant state, we have $\Phi_\sigma = 0$ for all $\sigma \in K$. Let us assume that

$$\hat{\mathbf{f}}^b_\sigma = \mathbf{f}(u^h)\cdot\mathbf{N}_\sigma \tag{28}$$

with $\sum_{\sigma\in K}\mathbf{N}_\sigma = 0$: this is the case for all the examples we consider. The flux $\mathbf{f}(u^h)$ has components on the canonical basis of $\mathbb{R}^d$: $\mathbf{f}(u^h) = \big(f_1(u^h),\dots,f_d(u^h)\big)$, so that

$$\hat{\mathbf{f}}^b_\sigma = \sum_{i=1}^d f_i(u^h)\mathbf{N}^i_\sigma.$$

Applying this to $\big(\hat{\mathbf{f}}^b_{\sigma_1},\dots,\hat{\mathbf{f}}^b_{\sigma_{\#K}}\big)$, we see that the j-th component of $\mathbf{n}_{\sigma,\sigma'}$ for $[\sigma,\sigma']$ direct, must satisfy:

$$\text{for any } \sigma\in K,\ \mathbf{N}^j_\sigma = \sum_{[\sigma,\sigma']\ \text{edge}} \varepsilon_{\sigma,\sigma'}\mathbf{n}^j_{\sigma,\sigma'}$$

i.e.

$$\big(\mathbf{N}^j_{\sigma_1},\dots,\mathbf{N}^j_{\sigma_{\#K}}\big)^T = A\ \big(\mathbf{n}^j_{\sigma,\sigma'}\big)_{[\sigma,\sigma']\in\mathcal{E}^+}.$$

We can solve the system and the solution, with some abuse of language, is

$$\big(\mathbf{n}_{\sigma,\sigma'}\big)_{[\sigma,\sigma']\in\mathcal{E}^+} = A^T L^{-1}\big(\mathbf{N}_{\sigma_1},\dots,\mathbf{N}_{\sigma_{\#K}}\big)^T \tag{29}$$

This also defines the control volumes since we know their normals. We can state:

Proposition 4.3. *If the residuals $(\Phi_\sigma)_{\sigma\in K}$ and the boundary fluxes $(\hat{\mathbf{f}}^b_\sigma)_{\sigma\in K}$ satisfy (25c), and if the boundary fluxes satisfy the consistency relations (28), then we can find a set of consistent flux $(\hat{\mathbf{f}}_{\sigma,\sigma'})_{[\sigma,\sigma']}$ satisfying (25). They are given by (27). In addition, for a constant state,*

$$\hat{\mathbf{f}}_{\sigma,\sigma'}(u^h) = \mathbf{f}(u^h)\cdot\mathbf{n}_{\sigma,\sigma'}$$

for the normals defined by (29).

We can state a couple of general remarks:

Remark 4.4.

1. *The flux $\hat{\mathbf{f}}_{\sigma,\sigma'}$ depend on the Ψ_σ and not directly on the $\hat{\mathbf{f}}^b_\sigma$. We can design the fluxes independently of the boundary flux, and their consistency directly comes from the consistency of the boundary fluxes.*
2. *The residuals depends on more than 2 arguments. For stabilized finite element methods, or the non linear stable residual distribution schemes, see e.g. [21, 16, 11], the residuals depend on all the states on K. Thus the formula (27) shows that the flux depends on more than two states in contrast to the 1D case. In the finite volume case however, the support of the flux function is generally larger than the three states of K, think for example of an ENO/WENO method, or a simpler MUSCL one.*
3. *The formula (27) are influenced by the form of the total residual (24). We show in the next paragraph how this can be generalized.*
4. *The formula (27) make no assumption on the approximation space V^h: they are valid for continuous and discontinuous approximations. The structure of the approximation space appears only in the total residual.*

4.2.2 Some particular cases: fully explicit formula

Let K be a fixed triangle. We are given a set of residues $\{\Phi^K_\sigma\}_{\sigma\in K}$, our aim here is to define a flux function such that relations similar to (23) hold true. We explicitly give the formula for $\mathbb{P}^1$ and $\mathbb{P}^2$ interpolant.

The general $\mathbb{P}^1$ case. The adjacent matrix is

$$A = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

A straightforward calculation shows that the matrix $L = A^T A$ has eigenvalues 0 and 3 with multiplicity 2 with eigenvectors

$$R = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix}$$

To solve $A\hat{\mathbf{f}} = \Psi$, we decompose Ψ on the eigenbasis:

$$\Psi = \alpha_2 R_2 + \alpha_3 R_3$$

where explicitly

$$\alpha_2 = \tfrac{1}{\sqrt{2}}\big(\Psi_1 - 2\Psi_2 + \Psi_3\big)$$

$$\alpha_3 = \sqrt{\tfrac{3}{2}}\big(\Psi_1 - \Psi_3\big)$$

so that

$$\hat{\mathbf{f}} = \frac{1}{3}\begin{pmatrix} \Psi_1 - \Psi_3 \\ \Psi_2 - \Psi_3 \\ \Psi_3 - \Psi_2 \end{pmatrix}.$$

In order to describe the control volumes, we first have to make precise the normals $\mathbf{n}_\sigma$ in that case. It is easy to see that in all the cases described above, we have

$$\mathbf{N}_\sigma = -\frac{\mathbf{n}_\sigma}{2}.$$

Then a short calculation shows that

$$\begin{pmatrix} \mathbf{n}_{12} \\ \mathbf{n}_{23} \\ \mathbf{n}_{31} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} \mathbf{n}_1 - \mathbf{n}_2 \\ \mathbf{n}_2 - \mathbf{n}_3 \\ \mathbf{n}_3 - \mathbf{n}_1 \end{pmatrix}.$$

Using elementary geometry of the triangle, we see that these are the normals of the elements of the dual mesh. For example, the normal $\mathbf{n}_{12}$ is the normal of PG, see figure 2.

Relying more on the geometrical interpretation (once we know the control volumes), we can recover the same formula by elementary calculations, see [5].

The general example of the $\mathbb{P}^2$ approximation. Using a similar method, we get (see figure 3 for some notations):

$$\begin{aligned}
\hat{\mathbf{f}}_{14} &= \frac{1}{12}(\Psi_1 - \Psi_4) + \frac{1}{36}(\Psi_6 - \Psi_5) + \frac{7}{36}(\Psi_1 - \Psi_2) + \frac{5}{36}(\Psi_3 - \Psi_1) \\
\hat{\mathbf{f}}_{16} &= \frac{1}{12}(\Psi_4 - \Psi_1) + \frac{5}{36}(\Psi_5 - \Psi_1) + \frac{7}{36}(\Psi_6 - \Psi_1) + \frac{1}{36}(\Psi_3 - \Psi_2) \\
\hat{\mathbf{f}}_{46} &= \frac{2}{9}(\Psi_2 - \Psi_6) + \frac{1}{9}(\Psi_3 - \Psi_5) \\
\hat{\mathbf{f}}_{54} &= \frac{2}{9}(\Psi_5 - \Psi_2) + \frac{1}{9}(\Psi_5 - \Psi_1) \\
\hat{\mathbf{f}}_{42} &= \frac{7}{36}(\Psi_2 - \Psi_3) + \frac{5}{36}(\Psi_1 - \Psi_3) + \frac{1}{12}(\Psi_6 - \Psi_3) + \frac{1}{36}(\Psi_5 - \Psi_4) \\
\hat{\mathbf{f}}_{25} &= \frac{1}{36}(\Psi_2 - \Psi_1) + \frac{5}{36}(\Psi_3 - \Psi_5) + \frac{7}{36}(\Psi_3 - \Psi_5) + \frac{1}{12}(\Psi_3 - \Psi_6) \\
\hat{\mathbf{f}}_{53} &= \frac{1}{36}(\Psi_1 - \Psi_6) + \frac{5}{36}(\Psi_3 - \Psi_5) + \frac{7}{36}(\Psi_4 - \Psi_5) + \frac{1}{12}(\Psi_2 - \Psi_5) \\
\hat{\mathbf{f}}_{63} &= \frac{1}{36}(\Psi_4 - \Psi_3) + \frac{5}{36}(\Psi_5 - \Psi_1) + \frac{7}{36}(\Psi_5 - \Psi_6) + \frac{1}{12}(\Psi_5 - \Psi_2) \\
\hat{\mathbf{f}}_{65} &= \frac{1}{9}(\Psi_1 - \Psi_3) + \frac{2}{9}(\Psi_6 - \Psi_4)
\end{aligned}$$

Then we choose the boundary flux:

$$\hat{\mathbf{f}}_\sigma^b = \int_{\partial K} \varphi_\sigma \mathbf{n}\, d\gamma$$

and get:

$$\mathbf{N}_l = -\frac{\mathbf{n}_l}{6} \quad \text{if } l = 1, 2, 3$$

$$\mathbf{N}_4 = \frac{\mathbf{n}_3}{3} \qquad \mathbf{N}_5 = \frac{\mathbf{n}_1}{3} \qquad \mathbf{N}_6 = \frac{\mathbf{n}_2}{3}$$

The normals are given by:

$$
\begin{aligned}
\mathbf{n}_{14} &= \frac{1}{12}(\mathbf{N}_1 - \mathbf{N}_4) + \frac{1}{36}(\mathbf{N}_6 - \mathbf{N}_5) + \frac{7}{36}(\mathbf{N}_1 - \mathbf{N}_2) + \frac{5}{36}(\mathbf{N}_3 - \mathbf{N}_1) \\
\mathbf{n}_{16} &= \frac{1}{12}(\mathbf{N}_4 - \mathbf{N}_1) + \frac{5}{36}(\mathbf{N}_5 - \mathbf{N}_1) + \frac{7}{36}(\mathbf{N}_6 - \mathbf{N}_1) + \frac{1}{36}(\mathbf{N}_3 - \mathbf{N}_2) \\
\mathbf{n}_{46} &= \frac{2}{9}(\mathbf{N}_2 - \mathbf{N}_6) + \frac{1}{9}(\mathbf{N}_3 - \mathbf{N}_5) \\
\mathbf{n}_{54} &= \frac{2}{9}(\mathbf{N}_5 - \mathbf{N}_2) + \frac{1}{9}(\mathbf{N}_5 - \mathbf{N}_1) \\
\mathbf{n}_{42} &= \frac{7}{36}(\mathbf{N}_2 - \mathbf{N}_3) + \frac{5}{36}(\mathbf{N}_1 - \mathbf{N}_3) + \frac{1}{12}(\mathbf{N}_6 - \mathbf{N}_3) + \frac{1}{36}(\mathbf{N}_5 - \mathbf{N}_4) \\
\mathbf{n}_{25} &= \frac{1}{36}(\mathbf{N}_2 - \mathbf{N}_1) + \frac{5}{36}(\mathbf{N}_3 - \mathbf{N}_5) + \frac{7}{36}(\mathbf{N}_3 - \mathbf{N}_5) + \frac{1}{12}(\mathbf{N}_3 - \mathbf{N}_6) \\
\mathbf{n}_{53} &= \frac{1}{36}(\mathbf{N}_1 - \mathbf{N}_6) + \frac{5}{36}(\mathbf{N}_3 - \mathbf{N}_5) + \frac{7}{36}(\mathbf{N}_4 - \mathbf{N}_5) + \frac{1}{12}(\mathbf{N}_2 - \mathbf{N}_5) \\
\mathbf{n}_{63} &= \frac{1}{36}(\mathbf{N}_4 - \mathbf{N}_3) + \frac{5}{36}(\mathbf{N}_5 - \mathbf{N}_1) + \frac{7}{36}(\mathbf{N}_5 - \mathbf{N}_6) + \frac{1}{12}(\mathbf{N}_5 - \mathbf{N}_2) \\
\mathbf{n}_{65} &= \frac{1}{9}(\mathbf{N}_1 - \mathbf{N}_3) + \frac{2}{9}(\mathbf{N}_6 - \mathbf{N}_4)
\end{aligned}
$$

There is not uniqueness, and it is possible to construct different solutions to the problem. In what follows, we show another possible construction. We consider the set-up defined by Figure 3. The triangle is split

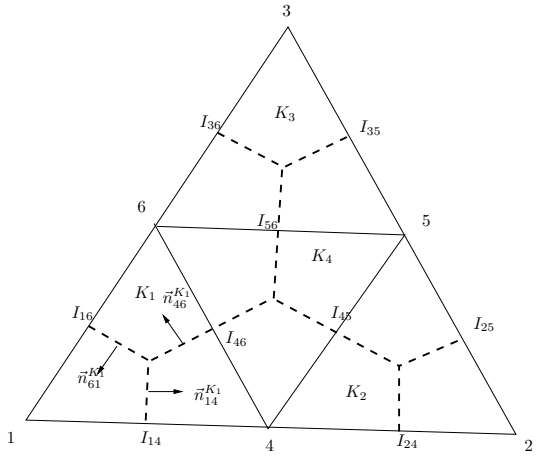

Figure 3: Geometrical elements for the $\mathbb{P}^2$ case. I_{ij} is the mid-point between the vertices i and j. The intersections of the dotted lines are the centroids of the sub-elements.

first into 4 sub-triangles K_1, K_2, K_3 and K_4. From this sub-triangulation, we can construct a dual mesh as in the $\mathbb{P}^1$ case and we have represented the 6 sub-zones that are the intersection of the dual control volumes and the triangle K. Our notations are as follow: given any sub-triangle K_ξ, if γ_{ij} is intersection between

two adjacent control volumes (associated to σ_i and σ_j vertices of K_ξ), the normal to γ_{ij} in the direction σ_i to σ_j is denoted by $\mathbf{n}_{ij}^\xi$. Similarly the flux across γ_{ij} is denoted $\hat{\mathbf{f}}_{ij}^\xi$.

Then we need to define boundary fluxes. If σ belongs to K_l, we denote the boundary flux as $\hat{\mathbf{f}}_\sigma^{K_l}$. A rather natural condition is that

$$\begin{array}{ll}\hat{\mathbf{f}}_l^{K_l} = \hat{\mathbf{f}}_l^b & l=1,2,3\\ \hat{\mathbf{f}}_4^{K_l} = \frac{1}{3}\hat{\mathbf{f}}_4^b & l=1,2,4\\ \hat{\mathbf{f}}_5^{K_l} = \frac{1}{3}\hat{\mathbf{f}}_5^b & l=2,3,4\\ \hat{\mathbf{f}}_6^{K_l} = \frac{1}{3}\hat{\mathbf{f}}_6^b & l=1,3,4.\end{array}$$

We recover the conservation relation. Other choices are possible since this one is arbitrary: the only true condition is that the sum of the boundary flux is equal to the sum of the $\hat{\mathbf{f}}_j^b$ for $j=1,\dots,6$: this is the conservation relation.

Then we set:

$$\begin{array}{lll}\Phi_1 &= -\hat{\mathbf{f}}^1_{\mathbf{n}_{61}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{14}} & +\hat{\mathbf{f}}^b_1\\ \Phi_2 &= -\hat{\mathbf{f}}^2_{\mathbf{n}_{42}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{25}} & +\hat{\mathbf{f}}^b_2\\ \Phi_3 &= -\hat{\mathbf{f}}^3_{\mathbf{n}_{53}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{36}} & +\hat{\mathbf{f}}^b_3\\ \Phi_4 &= -\hat{\mathbf{f}}^1_{\mathbf{n}_{14}} + \left(\hat{\mathbf{f}}^1_{\mathbf{n}_{46}} - \hat{\mathbf{f}}^4_{\mathbf{n}_{64}}\right) + \left(\hat{\mathbf{f}}^4_{\mathbf{n}_{45}} - \hat{\mathbf{f}}^2_{\mathbf{n}_{54}}\right) + \hat{\mathbf{f}}^2_{\mathbf{n}_{42}} & +\hat{\mathbf{f}}^b_4\\ \Phi_5 &= -\hat{\mathbf{f}}^2_{\mathbf{n}_{25}} + \left(\hat{\mathbf{f}}^2_{\mathbf{n}_{54}} - \hat{\mathbf{f}}^4_{\mathbf{n}_{45}}\right) + \left(\hat{\mathbf{f}}^4_{\mathbf{n}_{56}} - \hat{\mathbf{f}}^3_{\mathbf{n}_{65}}\right) + \hat{\mathbf{f}}^3_{\mathbf{n}_{53}} & +\hat{\mathbf{f}}^b_5\\ \Phi_6 &= -\hat{\mathbf{f}}^3_{\mathbf{n}_{36}} + \left(\hat{\mathbf{f}}^3_{\mathbf{n}_{65}} - \hat{\mathbf{f}}^4_{\mathbf{n}_{56}}\right) + \left(\hat{\mathbf{f}}^4_{\mathbf{n}_{64}} - \hat{\mathbf{f}}^1_{\mathbf{n}_{46}}\right) + \hat{\mathbf{f}}^1_{\mathbf{n}_{61}} & +\hat{\mathbf{f}}^b_6\end{array} \tag{30}$$

We can group the terms in (30) by sub-triangles, namely:

$$\begin{array}{lllll}\Phi_1 &= \left(-\hat{\mathbf{f}}^1_{\mathbf{n}_{61}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{14}} + \hat{\mathbf{f}}^b_1\right) & & &\\ \Phi_2 &= \left(-\hat{\mathbf{f}}^2_{\mathbf{n}_{42}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{25}} + \hat{\mathbf{f}}^b_2\right) & & &\\ \Phi_3 &= \left(-\hat{\mathbf{f}}^3_{\mathbf{n}_{53}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{36}} + \hat{\mathbf{f}}^b_3\right) & & &\\ \Phi_4 &= \left(-\hat{\mathbf{f}}^1_{\mathbf{n}_{14}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{46}} + \hat{\mathbf{f}}^{K_1}_4\right) & + & \left(-\hat{\mathbf{f}}^4_{\mathbf{n}_{64}} + \hat{\mathbf{f}}^4_{\mathbf{n}_{45}} + \hat{\mathbf{f}}^{K_4}_1\right) &\\ & & + & \left(-\hat{\mathbf{f}}^2_{\mathbf{n}_{54}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{42}} + \hat{\mathbf{f}}^{K_2}_4\right) &\\ \Phi_5 &= \left(-\hat{\mathbf{f}}^2_{\mathbf{n}_{25}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{54}} + \hat{\mathbf{f}}^{K_2}_5\right) & + & \left(-\hat{\mathbf{f}}^4_{\mathbf{n}_{45}} + \hat{\mathbf{f}}^4_{\mathbf{n}_{56}} + \hat{\mathbf{f}}^{K_4}_5\right) &\\ & & + & \left(-\hat{\mathbf{f}}^3_{\mathbf{n}_{65}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{53}} + \hat{\mathbf{f}}^{K_3}_5\right) &\\ \Phi_6 &= \left(-\hat{\mathbf{f}}^3_{\mathbf{n}_{36}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{65}} + \hat{\mathbf{f}}^{K_3}_6\right) & + & \left(-\hat{\mathbf{f}}^4_{\mathbf{n}_{56}} + \hat{\mathbf{f}}^4_{\mathbf{n}_{64}} + \hat{\mathbf{f}}^{K_4}_6\right) &\\ & & + & \left(-\hat{\mathbf{f}}^1_{\mathbf{n}_{46}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{61}} + \hat{\mathbf{f}}^{K_1}_6\right). &\end{array} \tag{31}$$

Then we define the sub-residuals per sub elements:

$$\begin{array}{ll}\Phi_1^1 = -\hat{\mathbf{f}}^1_{\mathbf{n}_{61}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{14}} + \hat{\mathbf{f}}^b_1\,, & \Phi_4^2 = -\hat{\mathbf{f}}^2_{\mathbf{n}_{54}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{42}} + \hat{\mathbf{f}}^{K_2}_4\\ \Phi_4^1 = -\hat{\mathbf{f}}^1_{\mathbf{n}_{14}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{46}} + \hat{\mathbf{f}}^{K_1}_4, & \Phi_2^2 = -\hat{\mathbf{f}}^2_{\mathbf{n}_{42}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{25}} + \hat{\mathbf{f}}^{K_2}_2\\ \Phi_6^1 = -\hat{\mathbf{f}}^1_{\mathbf{n}_{46}} + \hat{\mathbf{f}}^1_{\mathbf{n}_{61}} + \hat{\mathbf{f}}^{K_1}_6, & \Phi_5^2 = -\hat{\mathbf{f}}^2_{\mathbf{n}_{25}} + \hat{\mathbf{f}}^2_{\mathbf{n}_{54}} + \hat{\mathbf{f}}^{K_2}_5\\ & \\ \Phi_5^3 = -\hat{\mathbf{f}}^3_{\mathbf{n}_{65}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{53}} + \hat{\mathbf{f}}^{K_3}_5, & \Phi_4^4 = -\hat{\mathbf{f}}^4_{\mathbf{n}_{64}} + \hat{\mathbf{f}}^4_{\mathbf{n}_{45}} + \hat{\mathbf{f}}^{K_4}_4\\ \Phi_3^3 = -\hat{\mathbf{f}}^3_{\mathbf{n}_{36}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{65}} + \hat{\mathbf{f}}^{K_3}_3, & \Phi_5^4 = -\hat{\mathbf{f}}^4_{\mathbf{n}_{45}} + \hat{\mathbf{f}}^4_{\mathbf{n}_{56}} + \hat{\mathbf{f}}^{K_4}_5\\ \Phi_6^3 = -\hat{\mathbf{f}}^3_{\mathbf{n}_{36}} + \hat{\mathbf{f}}^3_{\mathbf{n}_{65}} + \hat{\mathbf{f}}^{K_3}_6, & \Phi_6^4 = -\hat{\mathbf{f}}^4_{\mathbf{n}_{56}} + \hat{\mathbf{f}}^4_{\mathbf{n}_{64}} + \hat{\mathbf{f}}^{K_4}_6,\end{array} \tag{32}$$

so we are back to the $\mathbb{P}^1$ case: in each sub-triangle, we can define flux that will depend on the 6 states of the element via the boundary flux. This is legitimate because in the $\mathbb{P}^1$ case, we have not used the fact that the interpolation is linear, we have only used the fact that we have 3 vertices. Clearly the fluxes are consistent in the sense of definition 4.1.

The same argument can be clearly extended to higher degree element, as well as to non triangular element: what is needed is to subdivide the element into sub-triangles.

The two solutions we have presented for the $\mathbb{P}^2$ case are different: the control volumes are different, since they have more sides in the second case than in the first one.

4.2.3 More specific examples

In what follow we look at the flux form on specific numerical schemes: an extension of the Rusanov scheme, what is called the N scheme after P.L. Roe and a discontinuous Galerkin method.

Rusanov residual. Here we assume a global continuous approximation. Assuming that the total residual is evaluated using the Lagrange interpolation of the flux, $\mathbf{f}^h = \sum_{\sigma' \in K} \mathbf{f}(u_\sigma)\varphi_\sigma$, we define (the integrals can be evaluated exactly in that case)

$$\Phi_\sigma(u^h) = \int_{\partial K} \varphi_\sigma \mathbf{f}^h \cdot \mathbf{n}\, d\gamma - \int_K \nabla\varphi_\sigma \cdot \mathbf{f}^h\, d\mathbf{x} + \alpha(u_\sigma - \bar{u}), \qquad \bar{u} = \frac{\sum_{\sigma' \in K} u_{\sigma'}}{\#K} \tag{33}$$

where $\#K$ is the number of degrees of freedom in K and α is a parameter that will become explicit later.

Since $0 = \int_K \varphi_\sigma \ \mathrm{div}\,(1)\, d\mathbf{x} = -\int_K \nabla\varphi_\sigma\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma \mathbf{n}\, d\gamma$ and $\sum_{\sigma' \in K} \varphi_{\sigma'} = 1$, we have

$$\begin{aligned}
\Phi_\sigma(u^h) &= \sum_{\sigma' \in K} \mathbf{f}(u_{\sigma'}) \cdot \Big(-\int_K \varphi_{\sigma'}\, \nabla\varphi_\sigma\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma\varphi_{\sigma'}\mathbf{n}\, d\gamma \Big) + \alpha(u_\sigma - \bar{u}) \\
&= \sum_{\sigma' \in K} \big(\mathbf{f}(u_{\sigma'}) - \mathbf{f}(u_\sigma)\big) \cdot \Big(-\int_K \varphi_{\sigma'}\, \nabla\varphi_\sigma\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma\varphi_{\sigma'}\mathbf{n}\, d\gamma \Big) + \alpha(u_\sigma - \bar{u}) \\
&= \sum_{\sigma' \in K} \Big(\big(\mathbf{f}(u_{\sigma'}) - \mathbf{f}(u_\sigma)\big) \cdot \Big(-\int_K \varphi_{\sigma'}\, \nabla\varphi_\sigma\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma\varphi_{\sigma'}\mathbf{n}\, d\gamma \Big) + \frac{\alpha}{\#K}\big(u_\sigma - u_{\sigma'}\big) \Big) \\
&= \sum_{\sigma' \in K} c_{\sigma\sigma'}(u_\sigma - u_{\sigma'})
\end{aligned}$$

with

$$c_{\sigma\sigma'} = -\frac{\mathbf{f}(u_\sigma) - \mathbf{f}(u_{\sigma'})}{u_\sigma - u_{\sigma'}} \cdot \Big(-\int_K \varphi_{\sigma'}\, \nabla\varphi_\sigma\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma\varphi_{\sigma'}\mathbf{n}\, d\gamma \Big) + \frac{\alpha}{\#K}.$$

A local maximum principle is obtained if for any element, and any couple of degrees of freedom in that element, we have $c_{\sigma\sigma'} \geq 0$. In the present case, we take

$$\alpha \geq \#K \max_{\sigma,\sigma' \in K} \left| \frac{\mathbf{f}(u_\sigma) - \mathbf{f}(u_{\sigma'})}{u_\sigma - u_{\sigma'}} \cdot \Big(-\int_K \varphi_{\sigma'}\, \nabla\varphi_\sigma\, d\mathbf{x} + \int_{\partial K} \varphi_\sigma\varphi_{\sigma'}\mathbf{n}\, d\gamma \Big) \right|.$$

In the case of triangular elements with $\mathbb{P}^1$ approximation, we have

$$\hat{\mathbf{f}}_{\sigma\sigma'} = \frac{1}{2}\Big(\int_K \nabla\big(\varphi_\sigma - \varphi_{\sigma'}\big) \cdot \mathbf{f}^h\, d\gamma \Big) + \alpha(u_\sigma - u_{\sigma'}).$$

Using simple geometry (see figure 2-a), we get

$$\hat{\mathbf{f}}_{\sigma\sigma'} = \frac{1}{|K|}\Big(\int_K \mathbf{f}^h\, d\mathbf{x} \Big) \cdot \mathbf{n}_{\sigma\sigma'} + \alpha(u_\sigma - u_{\sigma'}). \tag{34}$$

We see that this flux is not exactly the classical Rusanov flux

$$\hat{\mathbf{f}}^{Rus}_{\sigma\sigma'} = \frac{1}{2}\big(\mathbf{f}_\sigma + \mathbf{f}_{\sigma'}\big) \cdot \mathbf{n}_{\sigma\sigma'} + \alpha(u_\sigma - u_{\sigma'}),$$

but is formally very close to it: it is the sum of a centered part (the surface integral) and a dissipation. We also note that the flux (34) is not necessarily monotone, but it is monotone combined with the flux $\hat{\mathbf{f}}^b_{\mathbf{n}}$.

The N scheme. Considering the problem (1) with triangular elements. We assume the existence of an average vector $\overline{\nabla_u \mathbf{f}}$ such that

$$\frac{1}{2}\sum_\sigma \mathbf{f}_\sigma \cdot \mathbf{n}_\sigma = |K|\, \overline{\nabla_u \mathbf{f}} \cdot \nabla u^h.$$

Here, again both $\mathbf{f}$ and u are approximated by a linear Lagrange interpolant.

A simple example of such situation is given by the Burgers problem where $\mathbf{f}(u) = (\frac{u^2}{2}, u)^T$. Here

$$\overline{\nabla_u \mathbf{f}} = \big(\bar{u}, 1\big)^T$$

where $\bar{u}$ is the arithmetic average of the nodal values. This average is a generalisation of the Roe average [26], a version for the Euler equations can be found in [27].

Using this average, the N scheme, see [28], can be defined as follows:

$$\Phi_\sigma = k_\sigma^+\big(u_\sigma - \tilde{u}\big) \tag{35a}$$

with

$$k_\sigma = \frac{1}{|K|}\int_K \overline{\nabla_u \mathbf{f}} \cdot \nabla\varphi_\sigma \, d\mathbf{x}, \qquad k_\sigma^+ = \max(k_\sigma, 0), k_\sigma^- = \min(k_\sigma, 0) \tag{35b}$$

and

$$\tilde{u} = N\Big(\sum_{\sigma'\in K} k_{\sigma'}^- u_{\sigma'}\Big), \qquad N^{-1} = \sum_{\sigma'\in K} k_{\sigma'}^- \tag{35c}$$

The value of N is chosen such that the conservation (4) holds true. When looking at the flux $\hat{\mathbf{f}}_{\mathbf{n}_{\sigma,\sigma'}}$, no particular nice looking structure appears, except in the case of a thin triangle, where the associated flux is a generalisation of Roe's flux.

Note that Remark 3.1 also applies here, provided that the Rusanov residuals are replaced by those of the N scheme in the definition of β_σ, see (18).

Discontinuous Galerkin schemes ($\mathbb{P}^1$ case). The residual is simply

$$\Phi_\sigma^K = \oint_{\partial K} \varphi_\sigma \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma - \oint_K \nabla\varphi_\sigma \cdot \mathbf{f}(u^h)\, d\mathbf{x}.$$

In the $\mathbb{P}^1$ case, the flux between two DOFs σ and σ' is given by

$$\hat{\mathbf{f}}_{\sigma,\sigma'}(u^h, u^{h,-}) = \oint_{\partial K} (\varphi_\sigma - \varphi_{\sigma'})\hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma - \oint_K \nabla\big(\varphi_\sigma - \varphi_{\sigma'}\big) \cdot \mathbf{f}(u^h)\, d\mathbf{x}.$$

Again, from simple geometry,

$$\nabla\big(\varphi_\sigma - \varphi_{\sigma'}\big) = -\frac{\mathbf{n}_{\sigma\sigma'}}{|K|},$$

so that

$$\hat{\mathbf{f}}_{\sigma,\sigma'}(u^h, u^{h,-}) = \oint_{\partial K} (\varphi_\sigma - \varphi_{\sigma'})\hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma + \frac{\oint_K \mathbf{f}(u^h)\, d\mathbf{x}}{|K|} \cdot \mathbf{n}_{\sigma\sigma'}.$$

Note that $\oint_{\partial K}(\varphi_\sigma - \varphi_{\sigma'})\, d\gamma = 0$ if we take the same quadrature formula on each edge, as it is usually done. Hence, denoting by $\bar{u}$ the cell average of u^h on K, we can rewrite the flux as

$$\hat{\mathbf{f}}_{\sigma,\sigma'}(u^h, u^{h,-}) = \frac{\oint_K \mathbf{f}(u^h)\, d\mathbf{x}}{|K|} \cdot \mathbf{n}_{\sigma\sigma'} + \oint_{\partial K} (\varphi_\sigma - \varphi_{\sigma'})\big(\hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-}) - \mathbf{f}(\bar{u}) \cdot \mathbf{n}\big)\, d\gamma \tag{36}$$

so that the second term can be interpreted as a dissipation. The control volume is depicted in figure 4. Referring to figure 4 for the DOF #1, the flux on the faces $I1J$ is $\oint_{\partial K} \varphi_1 \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma$. In order to respect

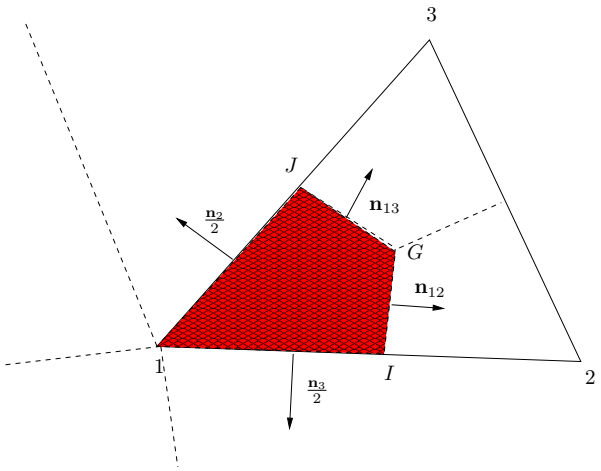

Figure 4: Representation of the control volume associated to DOF 1.

some geometrical assignment, the flux on $1I$ is set to

$$\oint_{1I} \varphi_1 \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma$$

and on $1J$,

$$\oint_{1J} \varphi_1 \hat{\mathbf{f}}_{\mathbf{n}}(u^h, u^{h,-})\, d\gamma.$$

5 Entropy dissipation

In this section, we consider the system version of (1). Our results on the flux are similar, since we never have used we were dealing with residual belonging to $\mathbb{R}$ or to some $\mathbb{R}^p$.

5.1 The 1 D case revisited

We start by recalling Tadmor's work [17, 18]. Let us start from a finite volume scheme semi-discretized in time:

$$\Delta x \frac{d\mathbf{v}_i}{dt} + \hat{\mathbf{f}}_{i+1/2} - \hat{\mathbf{f}}_{i-1/2} = 0.$$

If $\mathbf{v}$ is the entropy variable, we have:

$$\Delta x \langle \mathbf{v}_i, \frac{d\mathbf{u}_i}{dt} \rangle + \langle \mathbf{v}_i, \hat{\mathbf{f}}_{i+1/2} \rangle - \langle \mathbf{v}_i, \hat{\mathbf{f}}_{i-1/2} \rangle = 0$$

Then

$$\langle \mathbf{v}_i, \hat{\mathbf{f}}_{i+1/2} \rangle = \langle \frac{\mathbf{v}_i + \mathbf{v}_{i+1}}{2}, \hat{\mathbf{f}}_{i+1/2} \rangle + \langle \frac{\mathbf{v}_i - \mathbf{v}_{i+1}}{2}, \hat{\mathbf{f}}_{i+1/2} \rangle$$

Following Tadmor, we introduce the potential:

$$\theta = \langle \mathbf{v}, \mathbf{f} \rangle - \mathbf{g}$$

where $\mathbf{g}$ is the entropy flux, so that the entropy flux is defined by:

$$\hat{\mathbf{g}}_{i+1/2} := \langle \frac{\mathbf{v}_i + \mathbf{v}_{i+1}}{2}, \hat{\mathbf{f}}_{i+1/2} \rangle - \frac{\theta_i + \theta_{i+1}}{2}$$

and we get

$$\langle \mathbf{v}_i, \hat{\mathbf{f}}_{i+1/2} \rangle = \hat{\mathbf{g}}_{i+1/2} + \langle \frac{\mathbf{v}_i - \mathbf{v}_{i+1}}{2}, \hat{\mathbf{f}}_{i+1/2} \rangle - \frac{\theta_i + \theta_{i+1}}{2}$$
$$\langle \mathbf{v}_i, \hat{\mathbf{f}}_{i-1/2} \rangle = \hat{\mathbf{g}}_{i-1/2} + \langle \frac{\mathbf{v}_i - \mathbf{v}_{i-1}}{2}, \hat{\mathbf{f}}_{i-1/2} \rangle - \frac{\theta_i + \theta_{i-1}}{2}.$$

Thus,

$$\Delta x \langle \mathbf{v}_i, \frac{d\mathbf{u}_i}{dt} \rangle + \hat{\mathbf{g}}_{i+1/2} - \hat{\mathbf{g}}_{i-1/2} = \left(\langle \frac{\mathbf{v}_{i+1} - \mathbf{v}_i}{2}, \hat{\mathbf{f}}_{i+1/2} \rangle - \frac{\theta_{i+1} - \theta_i}{2} \right) + \left(\langle \frac{\mathbf{v}_i - \mathbf{v}_{i-1}}{2}, \hat{\mathbf{f}}_{i-1/2} \rangle - \frac{\theta_i - \theta_{i-1}}{2} \right).$$

This leads to the definition of entropy stable schemes:

Definition 5.1 (Tadmor [17, 18]). *A scheme is entropy dissipative if for any j,*

$$\langle \frac{\mathbf{v}_{j+1} - \mathbf{v}_j}{2}, \hat{\mathbf{f}}_{j+1/2} \rangle - \frac{\theta_{j+1} - \theta_j}{2} \leq 0$$

and entropy stable if we have an equality.

In residue form, we have

$$\Delta x \frac{d\mathbf{u}_i}{dt} + \phi_i^{i+1/2} + \phi_i^{i-1/2} = 0$$

with

$$\phi_i^{i+1/2} = \hat{\mathbf{f}}_{i+1/2} - \mathbf{f}_i, \qquad \phi_i^{i-1/2} = \mathbf{f}_i - \hat{\mathbf{f}}_{i-1/2}$$

so that for any j

$$\phi_j^{j+1/2} = \hat{\mathbf{f}}_{j+1/2} - \mathbf{f}_j, \qquad \phi_{j+1}^{j+1/2} = \mathbf{f}_{j+1} - \hat{\mathbf{f}}_{j+1/2}$$

If we compute $\langle \mathbf{v}_j, \phi_j^{j+1/2} \rangle + \langle \mathbf{v}_{j+1}, \phi_{j+1}^{j+1/2} \rangle$ (note this term is the one formulated in proposition 3.3), we get, using $\theta_j + \mathbf{g}_j = \langle \mathbf{v}_j, \mathbf{f}_j \rangle$

$$\begin{aligned}
\langle \mathbf{v}_j, \phi_j^{j+1/2} \rangle + \langle \mathbf{v}_{j+1}, \phi_{j+1}^{j+1/2} \rangle &= \langle \mathbf{v}_j, \hat{\mathbf{f}}_{j+1/2} - \mathbf{f}_j \rangle + \langle \mathbf{v}_{j+1}, \mathbf{f}_{j+1} - \hat{\mathbf{f}}_{j+1/2} \rangle \\
&= \langle \mathbf{v}_j - \mathbf{v}_{j+1}, \hat{\mathbf{f}}_{j+1/2} \rangle - \langle \mathbf{v}_j, \mathbf{f}_j \rangle + \langle \mathbf{v}_{j+1}, \mathbf{f}_{j+1} \rangle \\
&= \left(\langle \mathbf{v}_j - \mathbf{v}_{j+1}, \hat{\mathbf{f}}_{j+1/2} \rangle - \theta_j + \theta_{j+1} \right) + \mathbf{g}_{j+1} - \mathbf{g}_j.
\end{aligned}$$

So the condition

$$\langle \mathbf{v}_j, \phi_j^{j+1/2} \rangle + \langle \mathbf{v}_{j+1}, \phi_{j+1}^{j+1/2} \rangle \geq \mathbf{g}_{j+1} - \mathbf{g}_j$$

is equivalent to Tadmor's condition

$$\langle \frac{\mathbf{v}_{i+1} - \mathbf{v}_i}{2}, \hat{\mathbf{f}}_{i+1/2} \rangle - \frac{\theta_{i+1} - \theta_i}{2} \leq 0.$$

This suggests natural generalisation to the multidimensional case, i.e. the relation (20a).

5.2 The multidimensional case

Let us recall the entropy condition (20a),

$$\sum_{\sigma \in K} \langle \mathbf{v}_\sigma, \Phi_\sigma \rangle \geq \oint_{\partial K} \hat{\mathbf{g}}_{\mathbf{n}}(\mathbf{u}^h, \mathbf{u}^{h,-}) \, d\gamma.$$

Written like this, it seems that the residuals and the consistent entropy flux can be chosen independently, which is not exactly the case.

From the previous analysis, we have

$$\Phi_\sigma = \sum_{[\sigma,\sigma']} \hat{\mathbf{f}}_{\sigma\sigma'} + \hat{\mathbf{f}}^b_\sigma$$

with the condition (25c). This suggests to choose

$$\hat{\mathbf{f}}^b_\sigma = \oint_{\partial K} \varphi_\sigma \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}^h, \mathbf{u}^{h,-})\, d\gamma,$$

because (20a) becomes:

$$\sum_{\sigma\in K} \langle \mathbf{v}_\sigma, \Phi_\sigma\rangle = \oint_{\partial K} \langle \mathbf{v}^h, \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}^h, \mathbf{u}^{h,-})\rangle\, d\gamma + \sum_{\sigma\in K}\sum_{\sigma>\sigma'} \langle \mathbf{v}_\sigma - \mathbf{v}_{\sigma'}, \hat{\mathbf{f}}_{\sigma\sigma'}\rangle \geq \oint_{\partial K} \hat{\mathbf{g}}_{\mathbf{n}}(\mathbf{u}^h, \mathbf{u}^{h,-})\, d\gamma$$

Here we have set $\mathbf{v}^h = \sum_{\sigma\in K} \mathbf{v}_\sigma \varphi_\sigma$.

We introduce the potential θ^h in K by

$$\theta^h := \sum_{\sigma\in K} \theta_\sigma \varphi_\sigma \text{ with } \theta_\sigma = \langle \mathbf{v}_\sigma, \mathbf{f}(\mathbf{v}_\sigma)\rangle - \mathbf{g}(\mathbf{v}_\sigma). \tag{37}$$

Then we define $\hat{\mathbf{g}}_{\mathbf{n}}$ by

$$\hat{\mathbf{g}}_{\mathbf{n}}(\mathbf{u}^h, \mathbf{u}^{h,-}) = \langle \{\mathbf{v}^h\}, \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}^h, \mathbf{u}^{h,-})\rangle - \{\theta^h\}\cdot\mathbf{n}. \tag{38}$$

The numerical flux is defined only on ∂K and $\{a\}$ is the arithmetic average of the left and right states of a on the boundary of ∂K. The condition (20a) becomes

$$\sum_{\sigma>\sigma'} \langle \mathbf{v}_\sigma - \mathbf{v}_{\sigma'}, \hat{\mathbf{f}}_{\sigma\sigma'}\rangle + \oint_{\partial K} \theta^h_K \cdot \mathbf{n}\, d\gamma - \frac{1}{2}\left(\oint_{\partial K} \langle [\mathbf{v}^h], \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{v}^h, \mathbf{v}^{h,-}\rangle\, d\gamma - \oint_{\partial K} [\theta]\cdot\mathbf{n}\, d\gamma\right) \geq 0. \tag{39}$$

Here, the jump definition is consistent with Tadmor's definition in the one dimensional case: for any function w,

$$[w] = w_{|K^-} - w_{|K}. \tag{40}$$

From this we see that a sufficient condition for local entropy stability is that:

1. In K, we have

$$\sum_{\sigma\in K} \langle \mathbf{v}_\sigma, \Psi_\sigma\rangle + \oint_{\partial K} \theta^h_K \cdot \mathbf{n}\, d\gamma \geq 0, \tag{41a}$$

   where $\Psi_\sigma = \Phi_\sigma - \hat{\mathbf{f}}^b_\sigma$, or equivalently

$$\sum_{\sigma\in K}\sum_{\sigma>\sigma'} \langle \mathbf{v}_\sigma - \mathbf{v}_{\sigma'}, \hat{\mathbf{f}}_{\sigma\sigma'}\rangle + \oint_{\partial K} \theta^h_K \cdot \mathbf{n}\, d\gamma \geq 0. \tag{41b}$$

2. On the boundary of K we ask that the numerical flux $\hat{\mathbf{f}}$ is entropy stable so that

$$\oint_{\partial K} \left(\langle [\mathbf{v}^h], \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{v}^h, \mathbf{v}^{h,-})\rangle - [\theta]\cdot\mathbf{n}\right) d\gamma \leq 0. \tag{41c}$$

   Note that this condition is automatically met for a continuous $\mathbf{u}^h$.

The condition (41c) is automatically met is the flux $\hat{\mathbf{f}}$ is entropy stable in the sense of Tadmor:

$$\langle [\mathbf{v}^h], \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{v}^h, \mathbf{v}^{h,-})\rangle - [\theta^h]\cdot\mathbf{n} \leq 0. \tag{42}$$

Note that these conditions do not make any assumptions on the quadrature formulas on the boundary of K or in K. This is in contrast with the conditions on SAT-SBP schemes [3, 4, 1, 2].

Remark 5.2. *Starting from a consistent flux* $\hat{\mathbf{f}}$*, a simple way to construct a numerical flux* $\hat{\mathbf{f}}'$ *that satisfies* (42) *is to consider:*

$$\hat{\mathbf{f}}'_{\mathbf{n}}(\mathbf{v}^h, \mathbf{v}^{h,-}) = \hat{\mathbf{f}}_{\mathbf{n}}(\mathbf{v}^h, \mathbf{v}^{h,-}) + \alpha\big(\mathbf{v}^h - \mathbf{v}^{h,-}\big)$$

with α *chosen so that* (42) *holds true. If the original flux is Lipschitz continuous, this is always possible.*

Hence the satisfaction of (42) *is not an issue. Note this does not spoil the accuracy conditions* (22)*. Given a numerical flux, it is always possible to construct residuals that satisfies the conservation relation with that given flux. In the appendix, we show how to proceed for discontinuous representations. The next paragraph shows how to enforce a local entropy condition, in general.*

We can rework the relation (41b) in order to show some links with the recent paper [1]. Using the flux definitions, we can rewrite

$$\sum_{\sigma} \langle \mathbf{v}_\sigma, \Psi_\sigma \rangle + \oint_K \theta_{\mathbf{n}} d\gamma$$

as

$$\frac{1}{2} \sum_{\sigma > \sigma'} \big(\langle \mathbf{v}_\sigma - \mathbf{v}_{\sigma'}, \hat{\mathbf{f}}_{\sigma,\sigma'} \rangle - \big(\theta_\sigma - \theta_{\sigma'}\big) \cdot \mathbf{n}_{\sigma,\sigma'} \big) + \frac{1}{2} \sum_{\sigma > \sigma'} \big(\theta_\sigma - \theta_{\sigma'}\big) \cdot \mathbf{n}_{\sigma,\sigma'}$$

Then, we see that

$$\frac{1}{2} \sum_{\sigma > \sigma'} \big(\theta_\sigma - \theta_{\sigma'}\big) \cdot \mathbf{n}_{\sigma\sigma'} = \sum_{\sigma \in \partial K} \theta_\sigma \cdot \mathbf{N}_\sigma \, d\gamma = \oint_{\partial K} \theta_K^h \, d\gamma.$$

This relation is the motivation for defining θ^h in (37). Thanks to this, we can write the condition as:

$$\frac{1}{2} \sum_{\sigma > \sigma'} \big(\langle \mathbf{v}_\sigma - \mathbf{v}_{\sigma'}, \hat{\mathbf{f}}_{\sigma,\sigma'} \rangle - \big(\theta_\sigma - \theta_{\sigma'}\big) \cdot \mathbf{n}_{\sigma\sigma'} \big) + \oint_{\partial K} \theta(\mathbf{v}^h) \cdot \mathbf{n} \, d\gamma - \sum_{\sigma \in \partial K} \theta_\sigma \cdot \mathbf{N}_\sigma \geq 0.$$

with

$$\frac{1}{2} \sum_{\sigma > \sigma'} \big(\langle \mathbf{v}_\sigma - \mathbf{v}_{\sigma'}, \hat{\mathbf{f}}_{\sigma,\sigma'} \rangle - \big(\theta_\sigma - \theta_{\sigma'}\big) \cdot \mathbf{n}_{\sigma\sigma'} \big) \geq 0.$$

We see that, as in [1], if the fluxes $\hat{\mathbf{f}}_{\sigma,\sigma'}$ are entropy stable, we get entropy stability at the element level.

6 Conclusion

This paper shows some links between now classical schemes, such as the finite volume scheme, the continuous finite element methods, the discontinuous Galerkin methods and more generally a class of method nicknamed as Residual Distribution (RD) methods. We show that, under a proper definition of a consistent flux, all these schemes enjoy a flux formulation, and hence are locally conservative. This is well known for most schemes, less known for some of them. The fluxes are explicitly given. We also show that Tadmor's entropy stability condition can be reformulated very simply in the Residual Distribution context. Using this we have shown some connections with the recent work [1]. However the discussion here is certainly not finished, it will be the topic of another paper.

The emphasis of this paper is put on the steady case, but the unsteady state is similar, see [12] and [15].

Acknowledgements

The author has been funded in part by the SNSF project 200021_153604 "High fidelity simulation for compressible materials". I would also like to thanks Anne Burbeau (CEA-DEN) for her critical reading of the first draft of this paper. Her input has hopefully helped to improve the readability of this paper. The two referees and the editor are also warmly thanked for their patience, their comments and ability to trace typos. The remaining mistakes are mine.

References

[1] T. Chen and C.-W. Shu. Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws. *J. Comput. Phys.*, 345:427 – 461, 2017.

[2] A. Hiltebrand and S. Mishra. Entropy stable shock capturing spacetime discontinuous Galerkin schemes for systems of conservation laws. *Numer. Math.*, 126:103–151, 2014.

[3] G.J. Gassner. A skew-symmetric discontinuous Galerkin spectral element discretisation and its relation to SBP-SAT finite difference methods. *SIAM J. Sci. Comput.*, 35:A1233–A1253, 2013.

[4] J.E. Hicke, D.C. Del Rey, and D.W. Zingg. Multidimensional summation-by-part operators: general theory and application to simplex elements. *SIAM J. Sci. Comput.*, 38:A1935–A1958, 2016.

[5] R. Abgrall. On a class of high order schemes for hyperbolic problems. In *Proceedings of the International Conference of Mathematicians, volume IV*, pages 699–726, Seoul, 2014.

[6] R. Abgrall. Toward the ultimate conservative scheme: Following the quest. *J. Comput. Phys.*, 167(2):277–315, 2001.

[7] R. Abgrall. Essentially non-oscillatory residual distribution schemes for hyperbolic problems. *J. Comput. Phys.*, 214(2):773–808, 2006.

[8] R. Abgrall and P. L. Roe. High-order fluctuation schemes on triangular meshes. *J. Sci. Comput.*, 19(1-3):3–36, 2003.

[9] R. Abgrall. A residual method using discontinuous elements for the computation of possibly non smooth flows. *Adv. Appl. Math. Mech*, 2010.

[10] R. Abgrall and C.W. Shu. Development of residual distribution schemes for discontinuous Galerkin methods. *Commun. Comput. Phys.*, 5:376–390, 2009.

[11] R. Abgrall, A. Larat, and M. Ricchiuto. Construction of very high order residual distribution schemes for steady inviscid flow problems on hybrid unstructured meshes. *J. Comput. Phys.*, 230(11):4103–4136, 2011.

[12] M. Ricchiuto and R. Abgrall. Explicit Runge-Kutta residual distribution schemes for time dependent problems: second order case. *J. Comput. Phys.*, 229(16):5653–5691, 2010.

[13] R. Abgrall and D. de Santis. High-order preserving residual distribution schemes for advection-diffusion scalar problems on arbitrary grids. *SIAM J. Sci. Comput.*, 36(3):A955–A983, 2014. also http://hal.inria.fr/docs/00/76/11/59/PDF/8157.pdf.

[14] R. Abgrall and D. de Santis. Linear and non-linear high order accurate residual distribution schemes for the discretization of the steady compressible Navier-Stokes equations. *J. Comput. Phys.*, 283:329–359, 2015.

[15] R. Abgrall. High order schemes for hyperbolic problems using globally continuous approximation and avoiding mass matrices. *Journal of Scientific Computing*, 73:461–494, 2017.

[16] R. Struijs, H. Deconinck, and P.L. Roe. Fluctuation splitting schemes for the 2D Euler equations. VKI-LS 1991-01, 1991. Computational Fluid Dynamics.

[17] E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws, I. *Mathematics of Computation*, 49:91–103, 1987.

[18] E.Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 13:451–512, 2003.

[19] P. Ciarlet. *The finite element method for elliptic problems*. North-Holland, Amsterdam, 1978.

[20] A. Ern and J.L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer verlag, 2004.

[21] T.J.R. Hughes, L.P. Franca, and M. Mallet. A new finite element formulation for CFD: I. symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics. *Comp. Meth. Appl. Mech. Engrg.*, 54:223–234, 1986.

[22] E. Burman and P. Hansbo. Edge stabilization for Galerkin approximation of convection-diffusion-reaction problems. *Comput. Methods Appl. Mech. Engrg*, 193:1437–1453, 2004.

[23] D. Kröner, M. Rokyta, and M. Wierse. A Lax-Wendroff type theorem for upwind finite volume schemes in 2-d. *East-West J. Numer. math.*, 4(4):279–292, 1996.

[24] E. Burman, A. Quarteroni, and B. Stamm. Interior penalty continuous and discontinuous finite element approximations of hyperbolic equations. *J. Sci. Comput.*, 43(3):293–312, 2010.

[25] F. R.K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, 1997.

[26] P.L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43:357–372, 1981.

[27] H. Deconinck, P.L. Roe, and R. Struijs. A multidimensional generalization of Roe's flux difference splitter for the euler equations. *Computers and Fluids*, 22(2-3):215–222, May 1993.

[28] P.L. Roe and D. Sidilkover. Optimum positive linear schemes for advection in two and three dimensions. *SIAM J. Numer. Anal.*, 29(6):1542–1568, 1992.

A A DG RDS scheme

Let us consider problem (1) defined on $\Omega \subset \mathbb{R}^2$. In this case, the approximation can be discontinuous across edges: $u^h \in \mathcal{V}_h$.

In a first step, we consider a conformal triangulation of Ω using triangles. This is not essential but simplifies a bit the text. The 3D case can be dealt with in a similar way.

In K, we say that the degrees of freedom are located at the vertices, and we represent the approximated solution in K by the degree one interpolant polynomial at the vertices of K. Let us denote by u^h this piecewise linear approximation, that is in principle discontinuous at across edges. In the following, we use the notations described in Figure 5.

In [10], the degrees of freedom are located at the midpoint of the edges that connect the centroid of K and its vertices. This choice was motivated by the fact that the $\mathbb{P}^1$ basis functions associated to these nodes are orthogonal in $L^2(K)$. This property enables us to reinterpret the DG schemes as RD schemes, and hence to adapt the stabilization techniques of RD to DG. In particular, we are able to enforce a L^∞ stability property. However, this method was a bit complex, and it is not straightforward to generalize it to more general elements than triangles.

The geometrical idea behind the version that we describe now is to forget the RD interpretation of the DG scheme and to let the geometrical localization of the degrees of freedom move to the vertices of the element.

With this in mind, we define two types of total residuals:

- A total residual per element K

$$\Phi^K(u^h) = \oint_{\partial K} \mathbf{f}(u^h) \cdot \mathbf{n} \, d\gamma.$$

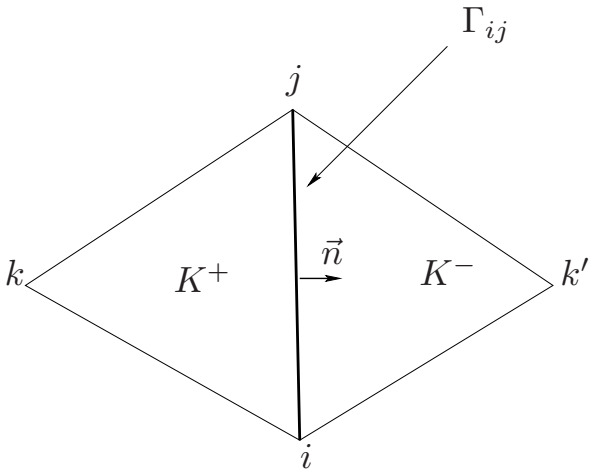

Figure 5: Geometrical elements for defining the scheme.

- A total residual per edge Γ, i.e.

$$\Phi_\Gamma(u^h) = \oint_\Gamma \left[\mathbf{f}(u)\cdot\mathbf{n}\right] d\gamma$$

where $[\mathbf{f}(u)\cdot\mathbf{n}]$ represents the jump of the function $\mathbf{f}(u)\cdot\mathbf{n}$ across Γ. Here, if $\mathbf{n}$ is the outward unit normal to K (see figure 5), which enables us to define a right side and a left side. Hence we set

$$[\mathbf{f}(u)\cdot\mathbf{n}] = (\mathbf{f}(u_R) - \mathbf{f}(u_L))\cdot\mathbf{n}.$$

We notice that Φ_Γ only depends on the values of u on each side of Γ.

The idea is to split the total residuals into sub-residuals so that a monotonicity preserving scheme can be defined. Here, we choose the Rusanov scheme, but other choices could be possible. Thus we consider

- For the element K and any vertex $\sigma \in K$,

$$\Phi_\sigma^K = \frac{\Phi^K}{3} + \alpha_K(u_\sigma - \overline{u}) \tag{43a}$$

with

$$\overline{u} = \frac{1}{3}\sum_{\sigma'\in K} u_{\sigma'},$$

and $\alpha_K \geq \max_{\mathbf{x}\in K} ||\nabla\mathbf{f}(u^h(\mathbf{x}))||$ where $||\ .\ ||$ is any norm in $\mathbb{R}^2$, for example the Euclidean norm.

- and for the edge Γ, any $\sigma \in \Gamma$,

$$\Phi_\sigma^\Gamma(u^h) = \frac{\Phi^\Gamma(u^h)}{4} + \alpha_\Gamma(u_\sigma - \overline{u}) \tag{43b}$$

with

$$\overline{u} = \frac{1}{4}\sum_{\sigma'\in K^+\cup K^-} u_{\sigma'}$$

where and $\alpha_\Gamma \geq \max_{K=K^+,K^-}\ \max_{\mathbf{x}\in\partial K\cap\Gamma} ||\nabla\mathbf{f}(u^h(\mathbf{x}))||$, see Figure 5 for a definition of $K^\pm$.

We have the following conservation relations

$$\begin{aligned}\sum_{\sigma\in K}\Phi_\sigma^K(u^h)&=\Phi^K(u^h),\\ \sum_{\sigma\in\Gamma}\Phi_\sigma^\Gamma(u^h)&=\Phi^\Gamma(u^h)\end{aligned}\tag{44}$$

The choice $\alpha_K\geq\max_{\mathbf{x}\in K}||\nabla\mathbf{f}(u^h(\mathbf{x}))||$ and $\alpha_\Gamma\geq\max_{K=K^+,K^-}\max_{\mathbf{x}\in\partial K\cup\Gamma}||\nabla_u\mathbf{f}(u^h(\mathbf{x}))||$ are justified by the following standard argument. If we set $Q=K$ or Γ, we can rewrite the two residuals as

$$\Phi_\sigma^Q(u^h)=\sum_{\sigma'\in Q}c_{\sigma\sigma'}^Q(u_\sigma-u_{\sigma'})$$

with $c_{\sigma\sigma'}^Q\geq 0$ under the above mentioned conditions. Indeed, using $u^h-u_\sigma=\sum_{\sigma'\in K}(u_\sigma-u_\sigma)\varphi_{\sigma'}$, we get (for $Q=K$ for example)

$$\begin{aligned}\Phi_\sigma^K(u^h)&=\frac{\Phi^K(u^h)}{3}+\alpha_K(u_\sigma-\overline{u})\\&=\frac{1}{3}\oint_{\partial K}\big(\mathbf{f}(u^h)-\mathbf{f}(u_\sigma)\big)\cdot\mathbf{n}\,d\gamma+\alpha_K(u_\sigma-\overline{u})\\&=\sum_{\sigma'\in K}\frac{1}{3}\left[\oint_{\partial K}\left(\int_0^1\nabla\mathbf{f}(su^h+(1-s)u_\sigma)\varphi_{\sigma'}(\mathbf{x})\,ds\right)\cdot\mathbf{n}\,d\gamma-\alpha_K\right](u_\sigma-u_{\sigma'})\end{aligned}$$

which proves the result.

Using standard arguments, as defining u^h as the limit of the solution of

$$u_\sigma^{n+1}=u_\sigma^n-\omega_\sigma\left(\sum_{K,\sigma\in K}\Phi_\sigma^K(u^{h,n})+\sum_{\Gamma,\sigma\in\Gamma}\Phi_\sigma^\Gamma(u^{h,n})\right)\tag{45}$$

with

$$\omega_\sigma\left(\sum_{K,\sigma\in K}c_{\sigma\sigma'}^K+\sum_{\Gamma,\sigma'\in\Gamma}c_{\sigma\sigma'}^\Gamma\right)\leq 1,$$

we see that we have a maximum principle.

It is possible to construct a scheme that is formally second order accurate by setting

$$\Phi_\sigma^{K,\star}(u^h)=\beta_\sigma^K\Phi^K(u^h)\text{ and }\Phi_\sigma^{\Gamma,\star}(u^h)=\beta_\sigma^\Gamma\Phi^K(u^h)\tag{46}$$

with

$$x_\sigma^K=\frac{\Phi_\sigma^K(u^h)}{\Phi^K(u^h)},\qquad x_\sigma^\Gamma=\frac{\Phi_\sigma^\Gamma(u^h)}{\Phi^\Gamma(u^h)},$$

and

$$\beta_\sigma^K=\frac{\max(x_\sigma^K,0)}{\sum_{\sigma'\in K}\max(x_{\sigma'}^K,0)},\qquad\beta_\sigma^\Gamma=\frac{\max(x_\sigma^\Gamma,0)}{\sum_{\sigma'\in K}\max(x_{\sigma'}^\Gamma,0)}.\tag{47}$$

As in the "classical" RD framework, the coefficients β are well defined thanks to the conservation relations (4). The scheme is written as (48) where the residuals $\Phi_\sigma^K(u^h)$ (resp. $\Phi_\sigma^\Gamma(u^h)$) are replaced by $\Phi_\sigma^{K,\star}(u^h)$ (resp. $\Phi_\sigma^{\Gamma,\star}(u^h)$.

The solution u^h is defined: find u^h linear in each triangle K such that for any degree of freedom σ (i.e. vertex of the triangulation),

$$\sum_{K,\sigma\in K}\Phi_\sigma^{K,\star}(u^h)+\sum_{\Gamma,\sigma\in\Gamma}\Phi_\sigma^{\Gamma,\star}(u^h)=0.\tag{48}$$

We have a first order approximation just by replacing the "starred" residuals by the first order ones. The system (48) is solved by an iterative method such as(45).